

What to do about Antarctica

Untrue to form, the members of the Antarctic Treaty have set out in good time to regulate mineral exploitation. Self-interest as well as fellow-feeling requires that they should succeed.

WHAT can possibly be the use of a continent covered in most places with ice, some of it several kilometres thick, which enjoys the worst climate in the world and which has no indigenous population? The simple answer seems to be that you can at least talk about the question. The Special Consultative Meeting of the members of the Antarctic Treaty that finished at Wellington in New Zealand two weeks ago (see *Nature* 10 February, p.457) has been doing just that. The meeting, the second in what will no doubt be a lengthy series (see below), had as its objective the definition of the circumstances under which the mineral resources of Antarctica might be exploited. By all accounts, the participating governments are no nearer a solution of the problem they have set themselves. Even the next meeting arranged for Bonn early in the summer is unlikely to bring a dramatic change. But now that the question has been opened, there are the strongest reasons why it should not be allowed to drop. For the consequences of that would jeopardize one of the most successful of international treaties signed since the Second World War.

As treaties go, the Antarctic Treaty is a remarkable instrument. Negotiated when the cold war of the 1950s was at its height, it has been most obviously successful in two respects: keeping Antarctica free from military enterprises, including nuclear weapons, and persuading the seven states with territorial claims to Antarctica to put them aside for at least the 30 years beginning with ratification in 1961. But the treaty has also worked well in an administrative sense. Scientific investigations have been mounted in Antarctica wherever and whenever those wishing to meet the cost have considered would be sensible. And some remarkable discoveries have been made, none more surprising than the discovery of fields of meteorites accumulated where the moving ice sheet is deflected by a mountain chain. The framework of the treaty has also served well as a means of safeguarding the environment. There is a convention on the exploitation of marine resources in the region (defined as that south of the sixtieth parallel) widely acknowledged to be imaginative and potentially effective. The members of the treaty like to boast how friendly are their discussions — the first meeting on mineral resources brought representatives of Britain and Argentina together more or less on the day on which Port Stanley in the Falkland Islands was reoccupied by the British Army.

Members' club

Part of the explanation seems to be that the treaty is a kind of club. There are twelve founder members, including all those with territorial claims, and now nearly as many others that have acceded to the treaty. The founder members have the special status of belonging automatically to consultative meetings such as those responsible for the convention on marine resources and, now, for negotiating a convention on minerals. Poland and West Germany have more recently joined this inner group, as permitted by the treaty, on the strength of their Antarctic research programmes. No doubt the recent interest of the Indian government in supporting expeditions to Antarctica is inspired by the calculation that if it should ever sign the treaty, it would be best to do so in such a way as promptly to be promoted to the inner council. Individuals participating at these meetings insist that the friendliness they recognize there is almost unprecedented in international negotiations intended to reconcile what are, in the last

resort, divergent political interests, and the record bears them out. But this should not be a surprise. The members of the inner club have the strongest reasons not to rock the boat, at least at this stage. For the members of the inner circle are effectively the collective managers of the continent. Thus they may amend the treaty by unanimous decision among themselves, whereupon other signatories of the treaty have no choice but to accept what has been agreed — or be deemed to have withdrawn. Better this, the members of the inner circle must be persuaded, than to have everything — territorial disputes included — put up for grabs.

How long this state of affairs will last is anybody's guess. The original treaty provides that any member of the consultative group may ask for a conference of all the parties to the treaty once 30 years have elapsed (which means 1991) to review progress or even to amend the treaty. This provision by itself provides the members of the consultative group with a further incentive to iron out differences among themselves. This, no doubt, is why the regulation of mineral exploitation has been taken up so far in advance of when Antarctic minerals, even off-shore oil, are likely to be economic, a few decades into the next century. (Sir Peter Scott seems to have been a little premature with his telex messages towards the end of the Wellington meeting complaining that Australia, France, Japan and West Germany were planning to drill for oil in the most inhospitable waters in the world.) Better that the club should agree now on what to do than be told by others how its members should behave.

Mineral exploitation

The obvious snag, the territorial claims of seven members of the consultative group, is an impediment but not an obstacle to progress. The terms of reference agreed at Buenos Aires in 1981 for the negotiations now under way require that the environmental impact of any proposals for mineral exploitation should be examined in advance, that the environmental integrity of the continent should take precedence over the economic advantages of exploitation and that whatever arrangements are made for approving schemes for mineral exploitation should be "open to states which commit themselves to objectives and principles of the Antarctic Treaty". Because one of the underlying principles is that "Antarctica shall continue forever to be used exclusively for peaceful purposes" even to the extent of an interdiction of "discord", and because the assertion of territorial rights entails at least the show of force, the participation of fourteen governments in the mineral negotiations is an implicit recognition that territorial claims will remain on ice so long as the treaty lasts. By the same token, however, it is safe to predict that territorial claims will not be voluntarily withdrawn; if the treaty should collapse, extant claims to territory would be more valuable than claims that had been abandoned. So the practical difficulty for the negotiators is to devise a means by which the exploitation of Antarctic minerals is supranationalized. The ideal would be that there should indefinitely coexist two quite separate ways of categorizing Antarctica by means of national identity, one based on the present flimsy territorial claims (likely to attenuate as time goes on), the other on the identities of those with viable and acceptable plans for extracting minerals (whose number may increase from zero).

How is such a trick to be accomplished? One thing is clear — the



members of the inner council of the Antarctic Treaty will not suddenly declare themselves for a version of the principle that has made a shambles of the Law of the Sea, the doctrine that a no man's land is by definition part of the "common heritage of mankind". The issue will be raised between now and 1991, but is not taken seriously. For, in principle at least, Antarctica is not a no man's land but a patchwork of often conflicting territorial claims. The notion that there should be some kind of equity in the exploitation of Antarctica between those who have explored some parts of it and those who have never been there would be especially offensive to the members of the consultative group now supporting the least effective research programmes. At a guess, the resentment of Argentina and Chile, both members of the consultative group with overlapping territorial claims, for the proposition that Uruguay (which is not) should have equal rights would be at least as great as that of governments that have spent handsomely on Antarctic research for a proposal that their freedom to win some benefit should be predicated on a demonstration that they have helped some other state, say Nicaragua or Romania, to follow in their footsteps.

The future of the treaty

The way forward is nevertheless clear and goes like this:

- First, having been told that environmental considerations are at some level paramount, the consultative group should set out to define the criteria that matter in proposals to look for oil on the southern continental shelf or for titanium in some outlandish outcrop. *Soi-disant* environmentalists should note that the survival of living things could be less important than the integrity of the West Antarctic ice-shelf, the huge partly grounded moving sheet of ice whose over-rapid melting might seriously inconvenience even people living in the Northern Hemisphere to such a point that their descendants would not exist. The members of the inner club should meet the cost, at least to begin with.

- Second, the group needs a mechanism for encouraging applications to exploit mineral deposits in Antarctica, if only to know what kinds of proposals to reject. The simplest course would be to invite proposals not from governments (with special axes to grind) nor from individuals (who in the last resort are less than infinitely substantial) but from corporate entities of its own devising. In the evolution of industrial society, the invention of the joint-stock company in the eighteenth century is acknowledged to have been crucial. Such an entity, while legally a person and free to enjoy the personal illusion that enterprise breeds success, was doubly accountable, both to its shareholders and to the law. The Antarctic Treaty's inner group, as hitherto represented by a mixture of diplomats and scientists, has been ill-equipped to invent the supranational equivalent of that device.

- Third (and the question will certainly be asked), what is to happen to the profits? The difficulty with Antarctica is that it is neither an obvious desert nor an obvious Eldorado but, for the time being, something in between. So even among the members of the treaty's inner group, united as they are in their inclination to remain exclusive, it is natural that everybody should dread the humiliation of learning that some fellow member has won a greater benefit. The first need, in such circumstances, is that the members of the management group should not be compelled to waste their own resources (on which there are other calls) on the pre-emptive exploitation of Antarctic minerals but that they should have a device for licensing their supranational versions of the joint-stock companies to operate for strictly limited periods of time (say ten years) under conditions so onerous that nobody can hope to become intolerably rich but only to make a return on capital employed substantially better than the international stock markets provide. But is that not merely postponing the problem? Yes, but not by a mere decade but by the best part of a century, by which time we shall all be wiser.

- Finally, however, could such a restrictive policy by the members of the Antarctic Treaty's inner group survive not merely their internal differences but the complaints of states still outside the treaty that a valuable international resource is needlessly being sterilized? Only time will tell. Complainants can appeal to the

International Court of Justice, whose decision in favour of a plaintiff would probably bring the treaty crashing down. The difficulty, for potential complainants, is that they can hope to win their legal cases only if they have carried out enough research to persuade others that they have justice on their side. But in that case, they might as well have acceded to the treaty in the first place, without going to The Hague. So as the decades pass, the pressure will grow to enlarge the inner group.

Lessons to learn

In the remote world of diplomacy, such developments will seem merely sensible accommodations between the uncertainties of the economic potential of Antarctica (and the unknown hazards of exploiting them) and the uncertain benefits that would flow from success. To professional people, the issue is more immediate, and for two reasons. First, the reason why the Antarctic Treaty emerged successfully like an unimagined phoenix from the cold war of a quarter of a century ago is that it was conceived of as a continuation by other means of what was known as the International Geophysical Year, a period of scientific collaboration not to be scorned because it lasted for only eighteen months. The preamble to the treaty makes its origins clear, and requires of signatory governments that they should give pride of place to the importance of collaboration with each other in the exploration of Antarctica. Second, the Antarctic Treaty has by accident become a precedent for other kinds of treaties. What would happen, for example, if the surface of some jovian satellite were found to be strewn with emeralds? Or if a sensible well-intentioned plan to convert solar energy into electrical power by means of Earth satellites were found to conflict with international law or even with the international treaty on the Peaceful Uses of Outer Space?

Compared with other continuing international negotiations, in places such as Beirut, Geneva, Madrid and Vienna, the negotiations on mineral exploitation in Antarctica are small beer. There is the best part of a decade before they have finally to be brought to a conclusion; with luck (the compliance of the members of the inner club), there may be even more time to play with. And there is no immediate danger that the managers of Antarctica will be forced to make a decision on some piratical plan to exploit mineral resources prematurely. In another sense, however, the Antarctic Treaty has by accident posed to those adhering to it an unprecedented question, that of how they will accommodate themselves to circumstances in which they have signed away many of the conventional attributes of national sovereignty. The states with territorial claims in Antarctica have accepted that there should be a moratorium, but they and all the others have accepted the conventionally even more demeaning rules that the activities of even publicly supported expeditions should be open to inspection by other signatories of the treaty, that expedition camps should not be fortified and that military personnel who find themselves in Antarctica should function not as soldiers but as scientists or explorers. Antarctica may be physically a wilderness, but diplomatically it is full of promise.

Risks to avoid

The fourteen Antarctic managers seem fully aware of these opportunities, as of the pitfalls that lie ahead. The danger of falling out among themselves is not great, while procedures for recruiting others to their number are laid down in the treaty as it exists. The most serious danger in the years immediately ahead, typified by what has been happening in Britain and in India, is that government may be persuaded by the importance of the Antarctic Treaty that no harm but only good can come from spending more each year on research in the Antarctic. In neither case is the extra sum involved so large that the general pattern of scientific activity is likely to be distorted. Yet the mere notion that more must be spent so that a "presence" should be more apparent is offensive. Perhaps the next step for the Antarctic Treaty's inner club should be to devise a test for linking membership with the quality of the research carried out and not simply with the cost of it. Or, better still, of setting up a grant-making agency financed by means of tax on what its members spend.

French science budget

Forecast 1983 expansion jeopardized by inflation

Sophia-Antipolis, France

THE swelling French science budget, voted at the end of 1982 to be 17.8 per cent greater in real terms this year than last, is in peril of being punctured by "rigueur" — the word French politicians use these days for "tightening the belt" or, more bluntly, for telling a minister that he will not get the funds he was promised.

There was rigueur last year too, but the Ministry of Research and Technology came out of it rather well, with an overall cut of around 10 per cent in current francs, according to latest estimates, leaving a substantial net real increase of 10 per cent. (The Ministry of Research and Technology's principal basic research council, the Centre National de la Recherche Scientifique (CNRS), was even better protected, seeing only a cut of FF50 million in a research budget, excluding salaries, of FF1,800 million.)

This year, however, the rigueur threatens to be harder. Prime Minister Pierre Mauroy, who must himself distribute the cuts to his ministers, is arguing strongly that if scientists are to be considered part of French society (they had complained of being forgotten), then they must also accept society's responsibilities. If rigueur turns into cuts of 25 per cent in current francs, the figure now being discussed, then science must take its share. Under such rigueur, the 17.8 per cent real increase in research budgets (plus 10 per cent inflation) would be nearly wiped out.

Thus Pierre Papon, director general of CNRS, has sent a letter to his laboratory directors forbidding them to spend more than 60 per cent of their promised and approved budget for 1983 — until they hear otherwise from him. Last year the cuts were restricted to large capital projects (dealt with by introducing slight delays) and "extra" money for certain research themes that CNRS hoped to promote; the ordinary running costs of laboratories were not affected. This year, as Papon's letters suggest and as Mauroy's argument implies, the money for materials and equipment may be affected also.

All this emerges as the achievements of the government in science and technology over the past year are put to the test in this little town in the south of France: Sophia-Antipolis, the ten-year-old science city on the Côte d'Azur. Four hundred scientists, administrators, technicians and industrialists have this week been asking whether the promises of the 3,000-strong national colloquium on science and technology in Paris last January have been

borne out. The answer so far is yes — with certain reservations. According to the organizer of this week's conference, biotechnologist Daniel Thomas of the University of Compiègne, the Paris meeting created a "psychological arena" among French scientists and industrialists:

Tasmanian dam

What price will buy off Hobart?

Canberra

THE tussle between the federal and state governments over the Tasmanian dam has been made more urgent by its general election, due on 5 March.

The Tasmanian government is considering a state cabinet submission valuing the resources of the south-west wilderness area at A\$2,500 million, a sum five times that offered by the Prime Minister, Mr Malcolm Fraser, on 20 January, by way of compensation for not building the Gordon-below-Franklin dam. Half of the A\$500 million package would cover the capital costs for building a coal-fired thermal power station of similar capacity to the proposed dam and the rest would subsidize the cost of the electricity generated to keep it on a par with hydro-electricity costs. By comparison, the dam will take 10 years to complete at a cost of A\$470 million and supply its output of 180 MW two to three times more cheaply.

Summarily dismissing the offer as being insufficient, Mr Robin Gray, the state premier, invited the federal government to make a higher bid. It appears that the state is planning not only more dams, but mines and timber logging on the world heritage site and the A\$2,500 million is allegedly the value of the total unexploited resources of the region.

Tasmania has been harder hit than the other states by the present recession and has the highest unemployment level. Industries are reluctant to move to Tasmania owing to the costs and delays of sea transport, emigration of skilled labour and the small domestic market of 400,000 people. One of the state's few attractions is its hydro-electricity, but cheap electricity alone cannot draw industries except those which have a large electricity bill, such as aluminium smelting.

Nevertheless the Tasmanian Hydro-Electric Commission (HEC) justifies the dam on the grounds that the present capacity will be exceeded not only by increased industrial development but also by increased domestic consumption.

they discovered for the first time since 1968 that they could talk to one another. At a pre-Sophia meeting in Lyons, last Friday, industry was pronouncing itself delighted at the number of doors that were opening to them on university research.

There are, however, still some complaints. The promised regional consultative committees on research have not been established, and certain regions such as the Rhône-Alpes would like much more autonomy. Also there is no sign, so far, of a common contract and career structure for all scientists, which would allow greater mobility among research organizations, universities and industry.

Robert Walgate

Another study by McLachlan Consultants, the group contracted by the Senate Select Committee on South-West Tasmania, using a different set of assumptions, disputes HEC's forecasts on demand, arguing that they are too high and that the dam or other alternatives need not be contemplated for another 3 years.

What appears certain, however, is that the dam will be a crucial issue in the general election. A recent poll shows 48 per cent of Australians are against the dam, 25 per cent are in favour of it and the rest are undecided. Its importance in turning the tide against the government is appreciated by some Liberal-National Party candidates in the seats at risk, who have declared themselves "anti-dammers". This is against the policy of their party, which decided against federal intervention to save the wilderness in December last year before it was placed on the world heritage list, and has since reiterated its position. The Australian Labor Party (ALP), on the other hand, which requires an additional 11 seats to take office, will, by a wide interpretation of the federal government's powers, legislate to stop the dam — a move which Tasmania may challenge in the High Court, the ultimate arbiter of constitutional power.

A coalition of conservation groups, including the Australian Conservation Foundation and the Tasmanian Wilderness Society, decided last week to campaign for ALP House of Representatives candidates in 13 mainland seats, presently held by the government, which are vulnerable to swing less than 2 per cent.

These considerations in a snap election have increased the pressure for a negotiated settlement over the dam issue. Such a solution would neither compromise state rights, nor lose the anti-dam vote on the mainland seats but would placate the pro-dammers in the five Tasmanian seats. But the precedent would be set for other states to demand compensation for not exploiting the resources in their world heritage sites.

Vimala Sarma

European Community research

Last goodbye to Super-Sara?

Brussels

ON the face of things, Britain could not have chosen a worse time to have led those urging the abandonment of the Super-Sara nuclear safety project at the research council held here last week. In the middle of the Sizewell inquiry and the debate on Britain's nuclear future, here was the United Kingdom anxious to cancel a major project conceived after the accident at Three Mile Island — an accident moreover which involved a pressurized water reactor (PWR) of the type which Britain plans to build at Sizewell.

On top of this, discussions on enhancing the importance of the Community's research activities are reaching a crucial stage and the United Kingdom is in general trying to dispel its partners' doubts as to whether the present government is positively interested in promoting a stronger Community.

Against all this, it must be said that Britain last week had the support of West Germany and the other member states except Greece and Italy in urging the project's abandonment, as well as an extremely negative experts' report. At the Research Council in December, Vicomte Etienne Davignon, the Commissioner for Industry, Energy and Research, had managed to postpone a final decision only by proposing that a trio of "wise men" should assess whether the project, given the escalation in costs, was still worth doing.

For Davignon, who nominated the experts himself, this seemingly adroit move proved, in the words of one British official, to be "an own goal". The report by Professors Birhofer, Holm and Teillac provided plenty of ammunition for the British and German delegations.

Although the report concedes that the project was originally a good idea, it concludes that because of the long delays, underestimates of budget and staff requirements and the possibility that the cost could continue to rise, the experiment was no longer cost-effective. Much the same results, it was felt, could be obtained from experiments being conducted elsewhere, and much sooner.

Despite the funeral atmosphere hanging over last week's council meeting, the last rites are still to be held over Super-Sara. The ministers will be meeting again on 10 March to consider what the alternatives are and the ministers took care to emphasize that neither Ispra nor the other three Joint Research Centres were at risk. The 293.9 million European Currency Units (ECU) (£161.7 million) set aside for the project would be spent instead on areas such as remote sensing, solar energy, fusion, the disposal of radioactive wastes or on the German suggestion of acid rain research. There was also some interest in turning the Joint Research Centres into

general European safety research centres.

Even so, much of the talk after the meeting was of apportioning the blame for a very expensive mess. So far, Super-Sara has cost more than 70 million ECU, and terminating the experiment, including the disposal of the Essor experimental reactor, will need up to a further 23 million ECU. And the United States will extract a substantial price for "European cooperation".

On Monday, the day before the Council meeting, Davignon had held a long press conference at which he roundly turned on the Council for procrastinating so long and causing the delays which had led to cost increases. The project was first approved in March 1980 but another year passed before the Council released a sum large enough (54 million ECU) to get going. Since then, doubts about the project have grown and led to postponement of other allocations.

The latest report points out that the conclusions of the "internal task force" released in April 1982 laid down a number of preconditions for success: that more money be immediately released, that the necessary personnel be assembled, that adequate management be appointed and that there should be no more delays. These conditions were never fulfilled.

This does not exonerate the Commission, however. The Commission gravely underestimated the cost of Super-Sara by a factor of 1.8. The management at Ispra has been extremely poor, while the Commission has failed to reform it in time or to present adequate alternatives.

The ten research ministers concluded from all this that "the decision-making processes for the preparation, adoption and implementation of the Joint Research Centres' programmes should be improved". But how? Davignon wants to have independent control of research programmes once they have been approved in principle and said that at the national level, ministers would not be found discussing research projects at in the detail customary in Brussels.

But the failure of Super-Sara reflects badly on nearly all concerned and must raise questions about the value of EEC research cooperation. Such misgivings were not, however, evident at this meeting. A favourable first reading was given to the three-year strategic framework programme which foresees raising research spending from 2.6 per cent of the EEC budget to 4 per cent. But there were few tangible suggestions as to how this might be done. "It was really a case of throwing out the one rotten tomato in a good bunch. Or, to mix metaphors, Super-Sara was the cuckoo in the nest that threatened the growth of other important research projects; it has to go", summarized one British official last week. **Jasper Becker**

US education

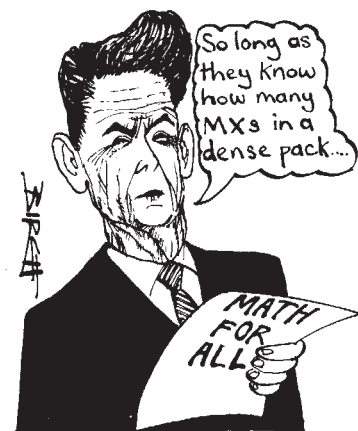
Money ahead

Washington

AMID new evidence of the worsening shortage of qualified mathematics and science teachers, the Reagan Administration last week presented to Congress the details of its plan for a crash programme to improve the teaching of mathematics and science in the nation's secondary schools.

The Administration's plan (filed in the House as HR 1324), which was first hinted at in the President's State of the Union address in January, calls for \$50 million to be spent each year on fellowships for up to 10,000 trainees. The money, which would be doled out through block grants to local school districts, would be aimed at teachers not currently in these subjects or would-be teachers who have science or mathematics backgrounds but not the necessary education courses. In either case, trainees would be chosen who would be ready to teach mathematics and science after just one year's training. The programme would continue for four years, training a total of 40,000 new teachers.

A \$20 million companion programme, administered by the National Science



Foundation (NSF), would seek to upgrade the training of an additional 10,000 teachers per year who are already teaching mathematics and science. This will probably be carried out through revitalizing the summer institutes that NSF has been running for teachers ever since the Sputnik launch set off the first wave of enthusiasm for science and mathematics improvement.

At a conference sponsored by the Department of Education last week, the proportions of the teacher shortage were outlined in vivid relief. According to data compiled by Betty Vetter, director of the non-governmental Scientific Manpower Commission, 22 per cent of high-school teaching posts in mathematics were vacant in 1982. In the Pacific coast states, where competition from the high-tech industries for technically-trained personnel is presumably especially high, 84 per cent of the

newly-employed science and mathematics teachers were not qualified to teach those subjects, she said.

"There is a growing body of evidence of diminishing quality as well", Vetter said. The average mathematics Scholastic Aptitude Test (SAT) score among all high school seniors last year was 467; among those seniors planning careers in education, it was 419. She added, "the number and academic standing of seniors planning to go into education were lower in 1980 than in 1972."

Paralleling this apparent decline in the competence of science and mathematics teachers is a continuing low level of performance in science and mathematics among students in general. Mathematics SAT scores have been declining for years, although they showed a slight upturn in 1982 (largely attributable to a dramatic rise in scores attained by minority students). Vetter noted that "half of all high school graduates take no math or science after the 10th grade". And the inevitable comparisons with other industrialized nations remain striking: in Japan, college-bound students take an average of three natural science courses and four mathematics courses in three years of high school. And, at college level, for every US engineering graduate, the United Kingdom turns out 1.1, West Germany 1.4 and Japan 2.6.

BAS's Falklands factor

THE British Antarctic Survey, conspicuous among British research organizations for having been offered more money to spend next year than it had hoped for, is planning that roughly £1 million of its extra £4 million should be spent in collaboration with universities. The survey's director, Dr R.M. Laws, says, however, that the recommendations of the survey's advisory committee have not yet been approved by the Natural Environment Research Council. One possibility is that there will be a seminar meeting with potential university collaborators in the early summer to decide which kinds of projects might be backed.

It is understood that the survey is also hoping with its own resources to embark on a marine geophysics programme, as well as to strengthen its work on marine living resources (based on South Georgia) and in particular its contributions to next year's international ice-age study and the coordinated study of marine resources planned for 1984-85.

Some of the windfall funds will be used to buy a trawler and aircraft, and to extend the airstrip at Rothera, on the west coast of the Antarctic Peninsula.

The extra £4 million available for 1983-84 is intended to increase to £5 million in succeeding years. In recent years, the survey's budget has not exceeded £6 million. Its unexpected good fortune seems to have been one of the uncovenanted benefits of the Falklands war. □

The difficulty of attracting trained scientists and mathematicians to teaching and keeping them there is exacerbated by a dismal salary picture. The average starting salary for 1982 graduates with bachelors' degrees taking jobs in industry or government was \$25,000 for engineers, \$21,000 for mathematicians, \$16,500 for biologists — and \$13,000 for teachers.

Some educators at last week's meeting doubted whether the Administration's proposal could do much given these structural problems. Others noted the absence of any funds for curriculum development, which had dried up since the days of the initial post-Sputnik wave of

enthusiasm. Dr George Keyworth, the President's science adviser, recently branded federal involvement in curriculum development as "social engineering".

A more ambitious proposal to revitalize science and mathematics teaching has already been approved by one House committee. This proposal, HR 1310, was introduced by Representative Carl Perkins (Democrat, Kentucky); it calls for the federal government to spend \$300 million in the first year. States could use these funds for teacher training, buying laboratory equipment, curriculum development or other programmes to improve science and mathematics teaching. **Stephen Budiansky**

Dead Sea project

Plan for filling sunken sea

ISRAEL'S long-discussed project to connect the Dead Sea with the Mediterranean has now reached the detailed planning stage and construction could be completed as early as 1990, according to the latest report from the Mediterranean Dead Sea Company. The main obstacle now is not so much the construction cost (estimated in January 1981 at US \$900 million) but opposition from the Hashemite Kingdom of Jordan, which regards the proposed construction as a breach of international law.

Politics apart, the case for the link-up is simple. Under natural conditions, the Dead Sea was a self-regulating hydrological system, with water-level maintained 393 m below sea level by an annual influx of some 1,600 million m³ of water from groundwater seasonal streams and, most importantly, the flow of the lower Jordan south from the Sea of Galilee. During the past two decades, however, the operation of the Israeli National Water Carrier, which draws its supply from the Sea of Galilee and the East Ghor Canal in Jordan, has reduced the inflow by 75 per cent. The result has been a 9 m drop in water level, the regression of the Dead Sea by up to several hundreds of metres and, following the 1979 drought, the actual splitting of the sea into two by a land-bridge formed at the Straits of Lisan.

Further hydroengineering works on the Jordanian side (including the damming of the Yarmuck and the extension of the East Ghor Canal) and increased demands for irrigation water on both sides of the border, are expected to increase the rate of shrinkage in the next decade. Without outside intervention the sea level is expected to fall to -408 m by 1990. The Mediterranean project aims to restore the water-level to the "historical" -393 m line, by bringing in water from the west. The cost of the project will be partly offset by generating hydroelectricity.

Since 1974, four possible routes for the Mediterranean link have been considered. That selected, on both economic and environmental grounds, is the most southerly, running from Qatif on the

Mediterranean, crossing the Negev south of Beer Sheva and entering the Dead Sea at Parsa, near the Straits of Lisan. The total route is some 110 km, of which 20 km will be an open canal following roughly the 100 m contour, with the crossing of the Judean mountains made by pressure tunnel. This system, it is planned, will carry 1,700 million m³ of water annually, until the Dead Sea is restored to the -393 m level, and thereafter 1,000 million m³ to maintain equilibrium. The generating capacity of the Parsa power station, at equilibrium, will be 1,080 million kwh per year.

The Jordanian objection to the plan is twofold. Firstly, Jordan claims that the proposed route is in breach of international law, since the intake installations are sited on "occupied Arab territory" — the Gaza Strip. Secondly, it says that the rise in sea level will result in the flooding of some of its best agricultural land, tourist hotels and roads, "antiquities" and the new Jordanian Potash Works, which started production last year and is expected to produce an annual income of \$200 million.

Part of the problem is that the Jordanians, as the water receded, built tourist facilities and the potash works right down to the water's edge whereas the Israelis, when the Dead Sea Works at S'dom became stranded inland, dug new channels to bring in brines to its evaporation beds.

Since the Israeli cost-effectiveness surveys include estimates for new construction at both the Dead Sea Works and the Jordanian Potash Works, it would appear that the Israelis might however meet reasonable compensation claims. The risk that total mixing of the Mediterranean and Dead Sea water would occur, making further extraction of potash impractical, would affect both plants equally, but is felt by the Israelis to be slight — an opinion not challenged by the Jordanians.

The Jordanian claim that good agricultural land and "antiquities" will be drowned seems less realistic, since the project is designed only to restore the water level to the pre-1960 mark. **Vera Rich**

Deep-sea drilling

Commercial break raises hopes

Washington

THE US deep-sea drilling community has just been saved — not by the bell but by the slump in world oil prices.

The community has been through years of cliff-hanging debate about which ship should serve for the next generation of drilling. The present vessel, *Glomar Challenger*, must be retired or overhauled at the end of the current season. The focus of this debate has been the giant *Glomar Explorer*, originally built for the Central Intelligence Agency. When the budget was made up last December, it seemed that the \$90 million cost of converting *Glomar Explorer* into a scientific ship was too high and that the plan would be scrapped.

But last week, a special advisory group headed by Dartmouth geophysicist Charles L. Drake told the National Science Foundation (NSF), which runs the programme, to forget *Explorer* because there were better ships on the horizon. It turns out that there are at least six highly capable commercial drillships now lying idle because the oil companies, which used to lease them to prospect for deep-water oil, are no longer doing so. The slump in oil prices has made deep-water exploration unattractive.

This is a boon to ocean science, the Drake group explained, because the government may be able to get an 8 to 10 year lease on one of these ships at a bargain price, and pay no more to convert it than it would have had to pay to modernize the old *Challenger*. Commercial drillships were ruled out during the *Explorer* debate because their running costs were twice those of either *Glomar Challenger* or *Explorer*.

The Drake group advised that oil prices may yet rise again, when industry may need the ships. So scientists have a "window" for obtaining a commercial ship at low rates. Sedco, a Texas company with two drillships, has even given NSF a public estimate of the costs and capabilities of one of their vessels converted for ocean drilling. The result is in the table below.

At the meeting with the Drake group, NSF assistant director Francis Johnson

nevertheless warned the deep-sea drillers of "the myth of budget isolation". In other words, he said, they could no longer consider their objective, a next-generation drillship, in isolation from other competing programmes in ocean geophysical science. The Drake group, which included oceanographers other than ocean drillers, had concluded that there is a need for continued drilling; it also concluded that financing the conversion of *Glomar Explorer* would impose too heavy a "tax" on other scientific programmes.

The commercial option has obviously excited many people in the deep-sea drilling community. The notion of doubling laboratory space, penetrating deeper into the ocean floor and getting a more stable, dynamically positioned ship that can stay on station for only \$10 million in conversion costs, is attractive. The real excitement, however, is that commercial drilling vessels come equipped with blow-out preventers — something *Glomar Explorer* might have had but only at extra cost. The continental margins are likely sites for hydrocarbon deposits. As director of the Scripps Institution of Oceanography William Nirenberg quips, now the ocean drillers can stop being "the only people avoiding the search for oil".

The ocean drillers also believe that a commercial vessel would increase the chance of foreign participation in the next phase of the programme. At present, the United Kingdom, France, West Germany, Japan and the Soviet Union contribute about one-third of *Glomar Challenger's* cost. But the Soviet Union has not been invited to participate in the next phase. The other foreign participants have said that the scientific returns from *Glomar Challenger* are diminishing.

The additional laboratory and berth space and the greater scientific potential of the commercial ship could make the programme more attractive to them. Indeed, US officials are considering making foreign participants pay half the costs of the programme, instead of one-third as at present.

Deborah Shapley

Possible US ocean drillships

Characteristic	<i>Glomar Challenger</i>	<i>Sedco 472</i>	<i>Glomar Explorer</i>
Length of drillpipe (ship to bottom of hole)	28,000 ft	30,000 ft	33,000 ft
Ship characteristics			
length	400 ft	470 ft	619 ft
beam	65 ft	70 ft	115 ft
depth	27 ft	32 ft	50 ft
displacement	10,600 t	16,664 t	40,000 t
Blowout preventer	No	Yes	No
High latitude capability	No	Possible	Yes
Laboratory space	4,500 sq.ft	9,000 sq.ft	19,000sq.ft
Daily operating cost (1983 dollars)	\$53,000	\$53,000	\$57,000
Cost of conversion (1983 dollars)	\$11,000,000	\$10,000,000	\$90,000,000

Source: National Science Foundation.

Soviet science writers

Medium seeks clearer message

SCIENCE journalists in the Soviet Union are not doing their job, according to a leading article in the Communist Party newspaper *Pravda*. Although many Soviet magazines and newspapers publish a steady flow of articles on scientific and technological achievements, the reader is rarely given a clear picture of what is going on, says *Pravda*. In particular, they are failing to pull their weight in what the 1981 Party Congress stressed as a major objective of Soviet science policy — the rapid implementation of research results.

Readers' letters, it seems, frequently complain of the "impersonal" presentation of science writing "decked out with numerical data", with little relevance to the working man.

A leading article in *Pravda* is a standard way of launching a new campaign or Party directive in the Soviet Union. *Pravda*, however, had few concrete suggestions to offer the journalists apart from a general hint that they should avoid stereotyped lists of achievements in favour of in-depth analysis.

Lists, however, appear to be an endemic feature of Soviet journalism. Last autumn, when the Moscow weekly *Literaturnaya Gazeta* decided to celebrate the diamond jubilee of the Soviet Union with reports on the state of science in the fourteen non-Russian Union Republics, it approached the presidents of the academies with three questions: what is the most interesting work your academy is doing, what are the main difficulties, in particular, as regards implementation of results in practice, and how is your academy participating in the All-Union Food programme?

The fact that the foremost literary weekly could fall into such a dullness trap — for the replies consisted in the main of fourteen near-identical lists — suggests that the style of science writing castigated by *Pravda* will not be easy to eradicate.

One participant in the survey, Habib Abdullaev of the Azerbaijan Academy, did seem to be aware of the pitfalls. He avoided giving a list of "achievements" altogether, and avoided the standard complaints of the lack of cooperation from bureaucracy and industry, a shortage of funds and the breakdown of essential supplies. Instead, he suggested that science journalism could help overcome one of the main problems of implementation — a lack of understanding between workers and scientists. Workers in industry and agriculture, he suggested, "should be told not only of the results of research, but also of the postulates and search which gave rise to them", if they are to respond appropriately to innovation. Whether the journalists can respond to the challenge remains to be seen.

Vera Rich

UK drug prescriptions

Mixed reaction to generics report

THE publication last week of a government-commissioned report recommending moves to encourage general practitioners (GPs) to prescribe more unbranded drugs has drawn howls of protest from the Association of British Pharmaceutical Industries (ABPI).

The report, by an Informal Working Group under the chairmanship of Dr P.R. Greenfield, was established to "identify ways of encouraging effective prescribing". Among its recommendations are more effective training of GPs in the cost effectiveness of alternative therapies and pleas for a more effective liaison between GPs and hospital doctors and it suggests that pharmacists dealing with GPs' prescriptions should be allowed to dispense alternative "generic" versions of an approved name formulation unless the doctor specifically indicates a preference for the proprietary brand. The report says that this could bring about "a significant change in the pattern of prescribing which would result in savings to the National Health Service".

Generic prescribing is already the practice in most hospital pharmacy departments, where local committees approve the generic substitutes used. In 1960 the Hindcliffe Commission recommended that all individual patients' prescriptions should use official titles, but by 1980 only 20 per cent of prescriptions were written by approved names, and the Greenfield report draws attention to the "commercial pres-

ures" exerted on doctors through advertising. It also notes that proprietary names, unlike chemical titles, are devised so as to be memorable.

The report comes at a sensitive time for the pharmaceutical industry, which was the subject of adverse comment following a recent BBC "Panorama" programme which dealt with drug companies' promotion practices. ABPI says that a move towards generic prescribing would cause an "immediate loss of confidence in the industry and result in less pharmaceutical research being located in Britain".

At present, according to ABPI, 10 per cent of world pharmaceutical research is carried out in the United Kingdom (although the market is only 4 per cent of the world total) and in 1982 it generated a £600 million balance of payments surplus. Generic substitution, it says, would result in price increases and a "substantial reduction in the present 70,000 workforce in the pharmaceutical industry itself and the peripheral supplying industries involving thousands more".

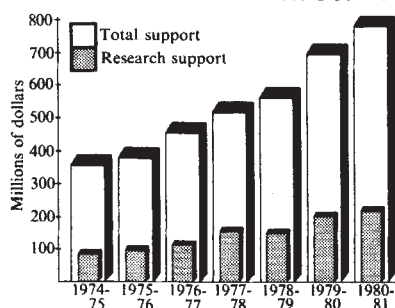
ABPI also questions the suggested savings to the National Health Service: Mr Peter Lumley, public affairs manager of ABPI, says that additional safety measures would have to be introduced to ensure that substitutes were subject to the same high quality control standards as the branded versions. The Pharmaceutical Society of Great Britain will cooperate with the Department of Health and Social Security to find ways of encouraging more generic prescribing but agrees on the need to ensure that substitutes are really identical, since under the new proposals patients would have no guarantee that they were getting the same preparation from one prescription to the next. The British Medical Association is unlikely to object to the moves in principle, but GPs will also need to be reassured about quality controls on unbranded drugs, and for some 30 or so common drugs no substitute is likely to be acceptable.

On the subject of GPs' training, the report recommends action to increase the awareness of cost effectiveness of different drugs, and even goes so far as to recommend compulsory courses in basic economic management techniques for all physicians. This recommendation was heartily endorsed by Dr Carne, senior tutor in General Practice at Hammersmith Medical School, who says that doctors seem to have forgotten about cost effectiveness in choosing therapies.

Mr Norman Fowler, Secretary of State for Social Services, has welcomed most of the report's findings but has balked at the prescribing suggestions on generic substitution. Until 15 April he is consulting on the matter with "interested parties".

Tim Beardsley

Industrial support of US academic research



CORPORATE support of academic research in the United States is greater than is generally believed, according to a report published by the National Science Board, the policy-making body of the National Science Foundation. The difference between the figures generally released and the actual figures is accounted for by university grants which are not reported as coming from industrial sources, by the donation or loan of equipment, and by the failure of universities to report grants from corporate sponsored foundations as being a type of industrial support. □

Paranormal hoaxer comes clean

St Louis

JAMES "the amazing" Randi held a press conference in late January to announce how he fooled researchers at Washington University who were studying the paranormal. Randi is a magician who has been associated with *Scientific American's* former columnist Martin Gardner, in a campaign to expose fraud in parapsychology.

Randi had sent two young self-proclaimed psychics, who were actually his protégés in the magic business, to the university's McDonnell Laboratory for Psychical Research in the autumn of 1979. Researchers at the laboratory studied them in 10 sessions between autumn 1979 and last July, observing their supposed ability to bend objects, produce images on film and read sealed letters. The laboratory published optimistic reports in the journal *Research on Parapsychology* in 1981 and 1982. The pair of psychics were the "most promising" subjects the laboratory had encountered, according to Peter Phillips, the director of the laboratory.

Randi played both sides of the experiment by also offering Phillips advice on



how to detect fraud in the psychics. In August 1981, Phillips sent a tape of the two psychics to Randi, who explained to him how they could be simply using magic. Phillips now admits "we should have taken his help earlier than we did". Also in August, Phillips's suspicions were aroused when he heard at a meeting that the two psychics might be fakes. In September he issued a statement saying the subjects were promising but that "ordinary explanations exist for these events".

Two more experiments were run with the pair, this time under more tightly controlled conditions, and Phillips became thoroughly suspicious. No more paranormal phenomena were observed in these tests, and the experimental programme was stopped.

The research was funded by the McDonnell Foundation and reflects the interest of James McDonnell, the founder of the giant McDonnell aircraft corporation, in psychic phenomena.

Phillips said he felt the laboratory came out looking good because his researchers had been cautious in their claims. And he endorses what Randi set out to achieve. "I don't resent Randi one bit — I think he's just great."

Karen Freeman

Animals in research

Congress gears up for legislation

Washington

BIOMEDICAL researchers were warned last week that a refusal to compromise on animal welfare legislation may play into the hands of militant animal rights groups that have vowed to increase their protests in the coming months.

The warning came from Mrs Christine Stevens, president of the Animal Welfare Institute, a moderate group that does not oppose the use of animals in research but which wants tougher rules to eliminate unnecessary pain and distress. Stevens appeared along with Dr Michael DeBakey, the Texas heart surgeon, at a forum on animal care sponsored by Howard University in Washington DC. Howard University was the victim of a Christmas-night break-in by a group calling itself the "Animal Liberation Front", which stole 28 cats and then announced its achievement to the *Washington Post*. The same group stole a dog from the Naval Medical Center in Bethesda two weeks later.

DeBakey held to his position that current safeguards are adequate to prevent abuses and that new legislation would "create more obstacles to doing research". Under the existing National Institutes of Health (NIH) rules, each institution receiving NIH research funds must have an animal care committee that monitors laboratory conditions and, at some institutions, reviews experimental protocols requiring the use of animals. "You can be sure that they're doing this as well as any federally-funded policemen", DeBakey said. "The mere fact that abuses take place in a small number of cases should not overshadow the fact that most scientists want to meet the guidelines set by NIH and the animal welfare groups."

Bills are expected to be reintroduced shortly in both the House of Representatives and the Senate, taking up where last year's failed efforts left off. The cornerstone of the legislation in the eyes of the Animal Welfare Institute is a requirement that the animal committees at NIH-supported institutions be required to take on one outside, non-scientist member who would specifically represent animal interests and that the committees be required (not just advised, as now) to review experimental protocols.

Both bills are likely to drop the provisions in last year's proposals that would have required accreditation of all federally-supported research facilities by a group such as the American Association for the Accreditation of Laboratory Animal Care. (Senator Robert Dole is drafting the Senate bill, Representative Doug Walgren the House bill.)

Stevens's plea for compromise comes at a time of increasingly militant action on the part of more extreme animal rights groups. An umbrella organization known as

Mobilization for Animals, which describes itself as a "direct action" group, is "targeting" four of NIH's seven regional primate centres for a protest on 24 April. Although Mobilization for Animals denies any involvement, several of the primate centres have reported attempted break-ins, apparently by animal rights groups.

The legislation may, however, be obviated by changes that NIH itself is considering in its regulations. According to Dr Robert Whitney of NIH, a draft of proposed changes "suggested a mandatory protocol review" by the institutional animal committees and the inclusion of an outside non-scientist. Whitney noted that NIH was also initiating an inspection pro-

gramme to determine how well the current system, which places the burden of responsibility for enforcement on institutions themselves, is working. The proposal has apparently been slowed down by some of the same opposition that DeBakey expressed with regard to the legislation, and is now being redrafted. Whitney emphasized that its final form has not yet been set.

DeBakey, in any event, said he was unimpressed by the argument that failing to support these changes would undercut the moderates, such as Stevens, and strengthen the extremists in the animal rights movement. "It's completely unrealistic to think that you're going to stop the antivivisectionists who want to stop all research in animals" with such legislation or new rules, he said.

Stephen Budiansky

Contraception

Djerassi calls for more research

New York

CARL DJERASSI, the Stanford chemistry professor who is known as the inventor of the contraceptive pill, continued his 15 year campaign for more research into birth control technologies last week in a memorial lecture at Rockefeller University in New York.

Djerassi decried the obstacles, which he called the "politics of contraception", standing in the way of control of the world's population growth. In his view, a range of technologies is needed, including devices for men and post-conception devices for women, in order to make birth control more acceptable in the developing world's varied societies.

One obstacle is the complacency caused by international family planning programmes in the 1970s, which lowered the rate of increase in the world's population. Djerassi is now, however, claiming that countries such as Mexico, which now has approximately 70 million people, could have 600 million people by 2050 if present trends continue.

Djerassi charged that under the Reagan Administration, the National Institutes of Health and the Agency for International Development have been cautious about funding research that might be construed as related to abortion.

Also, according to Djerassi, US manufacturers of oral contraceptives have been discouraged from developing new technologies because there are as many lawsuits filed involving oral contraceptives as there are for all other manufactured drugs combined.

Finally, the research programme Djerassi favours most, the World Health Organization's birth control programme, has been hurt by the continued refusal of the United States to give funds. The programme runs on approximately \$16 to \$20 million per year, contributed mostly by

Scandinavian countries. It is successful in part because it is international and avoids the stigma that a single nation encounters in selling so-called "genocidal" drugs to a poor country. But the foreign operations subcommittee of the Senate Appropriations Committee has consistently blocked US contributions.

Michael D. Stein

Polish civil rights

Physicist on trial

THE opening in Warsaw last week of the trial of Dr Zbigniew Romaszewski, physicist and civil rights activist, concided neatly with the resumption in Madrid of the review conference on the implementation of the Helsinki Final Act. For, although the formal charges against Dr Romaszewski deal with his alleged work for the clandestine "Radio Solidarity" (he was described in the court record as its "head and chief organizer"), the heavy sentence initially demanded by the prosecution — eight years — suggests that his other activities were tacitly taken into account.

In particular, Dr Romaszewski compiled, on behalf of the "Workers' Defence Committee" (KOR), the chief civil rights pressure group in Poland in the pre-Solidarity era, a report on breaches in Poland of the Helsinki provisions on human and civil rights, which was submitted to the Madrid review conference in November 1980. Like many KOR members, Dr Romaszewski was later closely involved with Solidarity, serving on its National Coordinating Commission, while his wife Zofia, also a physicist, and now indicted with him on the "Radio Solidarity" charge, served on the Solidarity Warsaw Region "Intervention Bureau" which dealt with complaints of unjust treatment and victimization of workers.

Vera Rich

Soviet psychiatry

Pre-emptive resignation?

THE Soviet Union last week resigned its membership of the World Psychiatric Association (WPA) because of Western pressure about the political misuse of psychiatry. The resignation was not entirely unexpected; Professor Marat Vartanyan, the Soviet representative on the WPA Ethics Committee, hinted last autumn that such a move was possible if Western national psychiatric associations continued with their plan to expel the Soviet All-Union Society of Neurologists and Psychiatrists (AUSNP) during the WPA General Assembly in Vienna next July.

At the previous WPA General Assembly (in Honolulu in 1977) a motion of censure against the Soviet Union was narrowly passed. Since then, several national psychiatric associations formerly opposed to sanctions against the Soviet Union have changed their attitudes, partly because of the further accumulation and dissemination of case histories by pressure groups such as the International Association on the Political Use of Psychiatry (IAPUP) but, according to Dr G. Bles, honorary secretary of IAPUP, chiefly because of the non-cooperative attitude of AUSNP.

The WPA General Assembly voted in 1977 to set up a special committee to monitor non-observance of the association's policy on the political use of psychiatry but AUSNP has consistently refused to acknowledge the committee's existence. This flouting of a decision of the General Assembly, said Dr Bles, probably helped sway many national societies.

By the end of last year, so many national societies had followed the initiative of the Royal College of Psychiatrists to expel AUSNP that the main problem in planning the agenda for Vienna was to decide which country should propose a motion to this effect. Now that the Soviet Union has offered what appears to be a pre-emptive resignation there seems little that the Western societies can do to bring pressure on their Soviet colleagues.

Sketchy details of two of the cases in which Soviet psychiatrists have allegedly confined dissident 'patients' on political rather than medical grounds were provided to the WPA executive committee last year. So far, however, none of the Soviet evidence has been sufficient to allay Western fears, and attempts to arrange for psychiatrists from the West to visit alleged 'prisoners of conscience' without unacceptable restrictions and conditions have met with no success. The official Soviet view, as propounded by Dr Aleksandr Snezhnevskii, is that all mental disturbance (including political dissent) is a type of schizophrenia.

The future WPA membership of other associations from the Soviet bloc remains in doubt. Further resignations, however, might be counter-productive. Although a few cases of the psychiatric confinement of dissidents have been reported in these countries they are, for the most part, isolated instances. A major walk-out would invite speculation that the practice is far more widespread than in fact appears to be the case.

Vera Rich

Princeton bid to strengthen life sciences

PRINCETON UNIVERSITY, obviously stung by recent complaints of its own departments' low standing in the life sciences, has announced that it will establish a new programme in molecular biology and raise \$46 million to that end. Princeton has hired Arnold J. Levine and Thomas E. Shenk, both at the State University of New York, Stony Brook, as the first members of what could grow to be a 19-faculty department by 1989. The university has also decided to raise \$29 million for a new building for the department to be set adjacent to the existing biological laboratory building.

Princeton as a whole ranked well in surveys of faculty quality made in 1969 and 1981, the second of which was released a year ago. But the ratings show that its life sciences departments have constantly slipped.

In 1969 Princeton's biochemistry department was ranked sixteenth in the United States but only twenty-second in 1981. In 1969 its programme in cellular/molecular biology ranked twelfth; in 1981

it ranked twenty-eighth. University officials are alarmed by the contrast with other departments ranked in the survey. Of the 26 Princeton departments ranked in the 1969 survey, only three were not in the first 16 — but two of the three low-rankers were biochemistry and cellular/molecular biology.

Princeton's zoology was ranked fifth in 1969 and not ranked in the latest survey. Other areas of biology, where the university is strong, such as population biology and neurobiology, were not ranked. Chemistry held thirteenth place in both surveys, but a recent advisory council report on the department urged that it must take account of molecular biology in order to keep pace.

Robert May, the Class of 1877 Professor of Zoology at Princeton, says that the university had been planning to improve its life science departments gradually, until the advisory councils for chemistry, biology and biochemistry weighed in with the opinions that drastic action was needed to remedy the situation. Deborah Shapley

US criminology

Chips has job, agenda follows

Washington

IN a classic movie, actor Jimmy Stewart played a supposed bumpkin thrust into the Washington scene. Now, in real life, one James Stewart, who likes to be called "Chips", a former chief of detectives for the city of Oakland, California, has been appointed by President Reagan to head the National Institute of Justice — the Justice Department's research agency that hands out approximately \$15 million a year for social science research in criminology and related fields.

University social scientists are described as "outraged" by the appointment of "Chips", who impresses them as well intentioned but not knowledgeable about criminal justice research. The institute, formerly part of the Law Enforcement Assistance Administration (LEAA), became a separate agency when LEAA was dismantled in 1980. Since its foundation in 1968, it has come to be a useful source of funds for the social sciences, especially when the National Science Foundation's social science budget declined to \$37 million per year.

Social scientists are even more concerned about the 17 new appointees to the institute's advisory board. They include two present or former policemen, a security guard from Las Vegas, three judges, two lawyers, two state legislators, a politician who is running the transition team of the new governor of California, a hotel executive from Pennsylvania and an advertising executive for Procter and Gamble. There is one academic in the new group, a law professor at the University of South Carolina.

The board has met four times, in Washington DC, New Orleans, Memphis and Atlanta, but has not put forward a programme for the current fiscal year. The aim of these hearings is to listen to practitioners so that the institute's research can be made more useful. The board has also expressed an interest in passing on individual grant applications — something research agency advisory boards rarely do.

Recipients of the institute's funds have also discovered that since becoming acting director last August (he was confirmed in December), Stewart has passed very few grant applications, with the result that the institute is now behind in its programme for fiscal years 1982 and 1983. The institute will, however, be able to spend the fiscal 1982 money once Stewart decides how it should be allocated.

In the movie, Jimmy Stewart triumphed in his Senate job and ended by doing a lot for the folks back home. Social science researchers, however, are still waiting to see what their James Stewart does for which of them.

Deborah Shapley

Seal hunting is no joking matter

SIR — A feature of the anti-sealing lobby that struck me from the outset (1965) as ominous was their use of the techniques so well described in *Animal Farm*. They have succeeded in persuading a very large number of people to see the issues in their own simplistic way. I find it incredible, however, that a scientific journal should print a political cartoon of the type seen in *Nature* of 23/30 December 1982 (p.674).

The early settlers of N.E. Newfoundland were faced with a severe continental climate bringing pack ice to the region for 5 months and making fishing impossible. Winters were passed in idleness until it was found that harp seals could be netted to provide oil, leather and meat, with the much later addition of fine furs. Alone among east coast Canadian fishing regions the inshore fishery of Newfoundland remains depressed due to excessive manpower and lack of alternative employment; sealing remains a valuable source of additional income. It is not generally understood that the winter inshore "fisheries" for harp seals off the Canadian east coast are diverse, taking various age categories between December and May. Thus a shift from a whitecoat hunt to the taking of other pelt types, if aesthetically desirable, would require no more than the intensification of an existing trend away from the use of large vessels to the seasonal use for sealing of medium-sized fishing vessels (longliners).

Further north in Labrador, a true arctic economy is practised by the Inuit or Eskimo (including Greenlanders) where sealing becomes the dominant source of income based on a renewable resource. The result of the propaganda aimed at eastern Canadian and other harp seal hunting has been to depress the prices of all harp seal pelts, such that the incomes of at least 100,000 people have been seriously jeopardized.

Part of the now popular simplistic approach to exploitation of seals and whales is to believe that once "victory" is achieved in legislating an end to hunting or buying, the problem is solved. However, the issues of sealing and whaling will always remain with us. First, the animals reproduce and grow slowly and so cannot be reared. Second, their pelagic populations do not compete with man for space, so that their habitats remain intact, as long as we maintain their food base, and the water quality of the near-seas. Consequently when hunting pressure is relaxed, sea mammal populations rise to near primeval levels, as we can now observe for many whale and seal populations around North America, and for grey seals around the United Kingdom. By contrast, wild ungulate populations on land — the closest parallel in terms of biomass — must be conserved also by safeguarding habitat or by domestication.

If therefore we are now passing through a period when whaling and sealing are socially unacceptable we ought to be working hard at two prerequisites of future exploitation: proper understanding of population dynamics and satisfactory methods of humane killing. Initiatives such as the Canadian government's proposal to examine North Atlantic sealing, together with the existing structure of the International Whaling Commission, provides a suitable framework for the first. Work on humane killing of seals and whales has been, by contrast, intermittent and under-financed.

D.E. SERGEANT

*Fisheries and Oceans Canada,
Arctic Biological Station,
Ste. Anne de Bellevue, Québec, Canada*

Botany leader

SIR — In the news item "US universities: Biological sciences ranked" (*Nature* 20 January, p.189) the statement is made that "In botany, the best were UC Davis and Vanderbilt". That should have been UC Davis and the University of Texas at Austin. While the error is understandable since it is easy to misread the listings, we worked too long and too hard to reach this ranking not to ask for a correction.

TOM J. MABRY

*Department of Botany,
University of Texas at Austin,
Austin, Texas, USA*

Stamp watching

SIR — In their interesting essay "Stamp watching and bird collecting" (ref.1) Robert and Judith May made one error so blatant that perhaps it was intentional — like their title. They expressed surprise that the number of bird species in North America and Britain "is almost exactly what would be expected from the MacArthur-Wilson species-area relation, $S \sim A^z$," because "the underlying premises from which the species-area relation is deduced pertain more to archipelagoes of islands than to large continental masses".

This misconception arises from a lack of historical perspective. The species-area relation is Arrhenius' law, $N = kA^z$, developed from botanical sampling². It seems applicable to most collections (even of stamps!) in both space and time and was extended in Preston's classic paper³ to the biota of continents and islands. MacArthur and Wilson⁴ contributed an hypothesis to explain the relationship as it applied to islands — the number of species being a balance between rates of migration and extinction — but even this is merely a

special case of the balance between species differentiation and extinction discussed by Preston for continents.

Indeed it is curious that "Many, if not most, of the subject's leading theoreticians in this century — Mayr, Lack, MacArthur and others — have been 'bird people' ". Preston is too.

JOHN E.C. FLUX

*Ecology Division, DSIR,
Lower Hutt, New Zealand*

1. May, R.M. & May, J. *Nature* 299, 11 (1982).
2. Preston, F.W. *Ecology* 41, 611-627 (1960).
3. Preston, F.W. *Ecology* 43, 185-215, 410-432 (1962).
4. MacArthur, R.H. & Wilson, E.O. *Evolution* 17, 373-387 (1963).

South Indian names

SIR — J.A. Irving (*Nature* 16 September 1982, p.200) states that Chinese surnames precede given names unlike Western conventions. This practice exists amongst the South Indians too. Like the Chinese, some of those who study in the West reverse this order while most don't, initializing the surname and the middle name. Thus S.N. Bose (of Bose-Einstein statistics), a Bengalee, follows the common Western and North Indian custom while C.V. Raman (of the Raman effect), a South Indian, has as his given name V. Raman and surname initialized to C.

The way the surnames are derived is different in the four South Indian states. The Andhras and the Kannadigas (of Andhra Pradesh and Karnataka) derive their family name through the paternal line, all generations having the same family name. The Tamilians (of Tamilnadu) have as their surname their father's given name with no constant family name. The Malayalees (of the state of Kerala) have two family names, one from the paternal line and the other from the maternal line, the males carrying the constant paternal family name and the females carrying the constant maternal name.

Since all computer forms now require names in the sequence last, first, middle perhaps the Western convention is on its way to change.

M.V. RAMANA

*Department of Physics,
West Virginia University,
Morgantown, West Virginia, USA*

Hopeful thinking

SIR — I do not wish to deter Mr Lawden from his desire to preserve the meaning of words. However, it is obvious from his letter (*Nature* 6 January, p.9) that it is too late to preserve the meaning of the word "hopefully".

D. L. SIMMS

*Department of the Environment,
London SW1, UK*

Sledgehammer cracks an apple

Two researchers at the University of Sussex reckon to have objective evidence that the siting of the 2.5m Isaac Newton Telescope in Britain was a "failure". But will anybody be surprised?

THE Isaac Newton Telescope has a long and chequered history, most of it as a concept rather than as an instrument. First endowed with the promise of government support in 1946, the telescope stands out among the world's major instruments in several unenviable ways. Design, construction and commissioning occupied no fewer than 21 years. For roughly ten years, the telescope was the only major optical telescope ever to have been operated at an altitude of only 30 m above sea-level and, the quality of the seeing apart, to have been so poorly sited that only 1,200 hours of usable telescope time were available each year. Moreover, the telescope is the largest ever to have been dismantled and moved lock, stock and barrel to another site, the new Northern Hemisphere Observatory on La Palma in the Canary Islands.

If that were not enough, John Irvine and Ben R. Martin from the Science Policy Research Unit at the University of Sussex have now applied to the brief operational history of the Isaac Newton Telescope what they describe as their method of "converging partial indicators" to argue that the telescope has made "a dismal showing" (*Social Studies of Science* 49, 86; 1983). Several questions arise, the least controversial being the unutility of the telescope.

By means of a painstaking comparison with similar telescopes elsewhere (especially the two-metre-plus instruments at Kitt Peak, Cerro Tololo and the Lick telescope on Mount Hamilton), Irvine and Martin show that the output of published scientific articles from the British instrument averaged only seven a year between 1969 and 1978, less than a fifth of the output from comparable telescopes elsewhere. Moreover, of the 71 scientific papers based on work carried out at the Isaac Newton Telescope during that decade, only two were cited as sources elsewhere more than 15 times in any year. Then, finally, 50 professional astronomers (including some from the Royal Greenwich Observatory where the telescope was sited) ranked the British instrument in a lowly place when asked to put the world's major telescopes in some kind of order of merit.

These are the partial indicators whose convergence points to the conclusion that the siting of this telescope on the edge of a stretch of Sussex marshland was, politely, a huge mistake. Perhaps the only comfort is that it might have been even worse if the Royal Greenwich Observatory had not been moved to Herstmonceux, in Sussex,

30 years ago. Then, London might have become the only metropolitan city in the world with a major telescope embedded in its street-lighting network. But, as things have turned out, nobody seriously dissents from the main conclusion that a major telescope should not have been sited on what was literally the sea-coast as recently as Roman times. Irvine and Martin might even openly have acknowledged that this perverse decision was condemned as such by many British astronomers from the outset and that it was one of the contentious issues dividing British astronomers throughout the 1960s.

So how valuable is it that this common view should now be lent an air of scholarly respectability? And what is the utility of the technique of converging partial indicators as a means of assessing decisions about large items of scientific equipment? Two other studies along the same lines by the same authors, dealing separately with British investments in high-energy physics and radioastronomy, are soon to appear in other journals and are understood to wag an equally reproving finger over past decisions. While, in the article now published, Irvine and Martin say that in the search for "partial indicators" that might have a bearing on the wisdom of decisions about new equipment or laboratories, the net should be cast as widely as possible, in practice they find themselves having to rely for what passes for statistical analysis on counts of published papers and citation analysis of the kind made possible by Dr Eugene Garfield of Philadelphia.

While nobody would pretend that data of this kind are irrelevant to an assessment of the productivity of a major instrument, the pitfalls in attempts at absolute measurement are also plain. Thus the small output of scientific papers based on the first year's operation of the Isaac Newton Telescope, numerically even less than accounted for by the poor climate, may be a consequence of the delay in carrying through the project: talented potential users may have gone elsewhere (many did), turned to radioastronomy or just given up hope. Indeed, the figures quoted by Irvine and Martin lend some support to this explanation — in the five years beginning in 1969, only 26 papers based on observations with the telescope appeared, compared with 45 in the succeeding half-decade, while, curiously enough, the frequency with which papers were cited elsewhere seems even from the outset to

have been respectable. In other words, the assembled data may also be taken as a sign that the telescope, for all its faults, helped powerfully in the resurgence of the now vigorous community of astronomical observers, instrumentalists and data handlers.

Unfortunately, the interviews conducted with astronomers in Britain and elsewhere yield only inconclusive evidence on these points. The opinions quoted of the telescope's reputation are not surprising. In other respects, the use made of the interviews falls awkwardly between the conventions of formal social science and journalism (British style). Thus several people are quoted anonymously as saying that ancillary equipment was inadequate, and the management of the Royal Greenwich Observatory insensitive to the needs of potential users from the universities, but Irvine and Martin do not offer their own opinion of the weight to be attached to these manifestly fair complaints. Nor does their argument throw much light on the reasons why this mistaken decision was made in the first place except to say that British astronomers were probably wrongly convinced that the government, in the decade after 1946, would not have supported a scheme to build a large telescope outside the United Kingdom.

The method of converging partial indicators is thus probably more useful as a means of confirming professional opinions about the foolishness of past decisions than of ensuring that future decisions will be wise. In 1946, Britain was still the centre of a huge empire. It was as natural that the government of the day should sanction the building of a large telescope as that it should decide, at the same time, to build nuclear weapons. The unforgivable error was that neither the astronomical community nor the British Admiralty (then the paymaster of the Royal Greenwich Observatory) sought to unravel the decision when it became plain, soon afterwards, that radioastronomy offered better opportunities. Yet now, after three decades of frustration and some expense, the old objective of revivifying observational astronomy seems to have been achieved. The value of international collaboration and of intelligent management in the interests of all potential users is also slowly being appreciated. But the question whether the game has been worth the candle is still more easily answered by hunch than by calculation. □

Growth hormone-releasing factor

A tale of two islets

from George Fink

THE two recent reports^{1,2} of the isolation and sequencing of human pancreatic growth hormone-releasing factor (hpGRF) from the laboratories of Roger Guillemin and Wylie Vale have rapidly been followed by immunohistochemical evidence from Guillemin's group that hpGRF is the same as the hypothalamic factor that normally stimulates growth hormone release. In this issue of *Nature* (p.607), Bloch and colleagues³ show that an antiserum against the hpGRF peptide specifically stains neural cell bodies in areas of the primate hypothalamus associated with the hypothalamic peptide (GRF).

The isolation and sequencing of hpGRF was only made possible by the existence of two islet cell tumours, one in a patient at Charlottesville, Virginia and another in a patient at Lyon. Carcinomas of the pancreatic islets (the islets of Langerhans) are relatively infrequent (less than 1/100,000 of the population⁴) and because they often secrete large amounts of different types of peptide hormones they are extremely difficult to diagnose. The tumours most commonly secrete the hormones indigenous to the pancreas — insulin (symptoms of hypoglycaemia), gastrin (gastric or duodenal ulceration and diarrhoea; Zollinger–Ellison syndrome), glucagon (mild diabetes mellitus and dermatitis) and somatostatin (hypochlorhydria, steatorrhoea and diabetes mellitus)^{4,5}. Less frequently they secrete hormones ectopic to the pancreas such as secretin (leading to 'pancreatic cholera'), adrenocorticotropin (presenting with features of Cushing's syndrome), serotonin (associated with watery diarrhoea, flushes and tachycardia)⁴ and in the present case, hpGRF (associated with acromegaly).

The isolation of hpGRF was carried out quite separately by two teams at the same institute — the Salk Institute — and headed respectively by Wylie Vale (Peptide Biology Laboratory) and his former mentor Roger Guillemin (Laboratories for Neuroendocrinology). Peptide isolation, like the Americas Cup, offers no second place, but in this case the race appears to have been a dead heat.

Vale's tumour was removed from a 21-year-old woman who presented at Mike Thorner's (formerly of Bart's) clinic in Charlottesville with Turner's syndrome and the signs and symptoms of acromegaly⁶. Guillemin obtained his tumour by way of the 'French connection'; it was removed from a man who presented with acromegaly at the clinic of G. Sassolas at Lyon¹. Both teams used essentially the same methods — gel filtration, HPLC and Edman degradation. Guillemin¹ isolated three peptides, the largest of which was 44

residues long and had the sequence H-Tyr-Ala-Asp-Ala-Ile-Phe-Thr-Asn-Ser-Tyr-Arg-Lys-Val-Leu-Gly-Gln-Leu-Ser-Ala-Arg-Lys-Leu-Leu-Gln-Asp-Ile-Met-Ser-Arg-Gln-Gln-Gly-Glu-Ser-Asn-Gln-Glu-Arg-Gly-Ala-Arg-Ala-Arg-Leu-NH₂. The peptide isolated by Vale² was 40 amino acid residues long and had a sequence identical to the first 40 residues of Guillemin's peptide. The two other factors found by Guillemin to release growth hormone comprised the 1–37 and 1–40 sequences of the 44-residue hpGRF.

Both Vale and Guillemin agree that the N-terminus is essential for biological activity [deletion of -Tyr (ref. 4), or Tyr-Ala or Tyr-Ala-Asp (ref. 1) leads to almost total loss of activity] but whereas Guillemin found that hpGRF_{1–37} had only 12 per cent of the activity of hpGRF_{1–44}, Vale² reported that the first 29 residues were sufficient for full expression of activity *in vitro*. The peptide has several sequence homologies with peptides of the glucagon-secretin family, but none of these peptides has growth hormone-releasing activity. The greatest homology is with a 27-residue peptide recently isolated from porcine gut, PHI-27 (ref. 7), which has 12 residues that are also present in equivalent positions in hpGRF_{1–27}.

The isolation of hypothalamic GRF, like the other hypothalamic regulatory peptides, has been a will-o-the-wisp. Andrew Schally caused a flurry of interest when in 1971 he published the structure of GRF (from porcine hypothalamus)⁸ as H-Val-His-Leu-Ser-Ala-Glu-Glu-Lys-Glu-Ala-OH, but although his peptide appeared to release growth hormone as assessed by bioassays (which used, in the main, the capricious pituitary growth hormone depletion–tibia assay), it was never found to release growth hormone as measured by radioimmunoassay.

Guillemin attributes the difficulty of isolating GRF from hypothalamus to the incredibly small tissue concentration in comparison with that of other hypothalamic neuropeptides (femtomoles compared with pico- and nanomoles). Guillemin and Vale both argued that hpGRF is likely to be identical to the hypothalamic GRF by analogy with other ectopic hormones, the structure of which has always been similar or identical to the structure of the same hormone produced by the tissue to which it is indigenous. Further, both groups had shown that partially purified rat hypothalamic GRF (rGRF) co-eluted with hpGRF on chromatography and that in terms of growth hormone released, the dose–response curves for rGRF were parallel to those for hpGRF.

The work reported from Guillemin's group in this issue³ (p.607) has come close to clinching the deal by showing that in the monkey and human, nerve fibres around the primary capillary plexus of the hypophysial portal vessels stain immunohistochemically with an antiserum raised in rabbits against hpGRF_{1–40}. Cell bodies in the tubular nuclei were also immunoreactive. The antiserum did not stain fibres or cell bodies in the hypothalamus of either the rat or the ox, perhaps not surprising since the antiserum was raised against a human peptide. Bloch *et al.*³ speculate that this may be due to species differences in the 28–40 region of the molecule which is devoid of GRF bioactivity. Species differences may also explain some of the differences in the GRF activity of extracts of porcine stalk median eminence compared with that of GRF partially purified from islet tumour tissue⁹.

Peptide isolation aside, there is much fundamental biology to be learned from the studies carried out on Thorner's patient⁶ as well as from the elegant case reports of Lawrence Frohman¹⁰ whose work on extrapituitary GRF activity must be considered as the stimulus and model for the studies by Guillemin and Vale. Thorner's patient showed the hallmarks of acromegaly due to a pituitary adenoma: plasma growth hormone concentrations were increased in response to glucose (instead of suppressed), increased after TRH administration (instead of no change) and decreased with dopamine infusion (instead of increased). Were these abnormal responses due to intrinsic changes in the hypertrophic pituitary somatotrophs or changes in ligand–receptor interactions brought about by excessive exposure to GRF, or did these factors (glucose, TRH and dopamine) act on the tumour directly to affect hpGRF release? What happens to somatostatin (in brain and pancreas) in these situations in view of the fact that somatostatin inhibits the growth hormone response to GRF release in a non-competitive manner^{5,9}? Whatever the answers to these and other questions regarding GRF biosynthesis and its role in gut and pancreas, the diagnosis and removal of two islet cell carcinomas has made two patients, two peptide chemists and several biological firms very happy. □

George Fink is Director of the MRC Brain Metabolism Unit, University Department of Pharmacology, 1 George Square, Edinburgh EH8 9JZ.

- Guillemin, R. *et al.* *Science* **218**, 585 (1982).
- Rivier, J., Spiess, J., Thorner, M. & Vale, W. *Nature* **300**, 276 (1982).
- Bloch, B. *et al.* *Biochem. biophys. Res. Commun.* **45**, 235 (1971).
- Schein, P.S., *et al.* *Ann. int. Med.* **79**, 239 (1973).
- Larsson, L.-I. *et al.* *Lancet* **i**, 666 (1977).
- Thorner, M.O. *et al.* *J. clin. Invest.* **70**, 965 (1982).
- Tatemoto, K. & Mutt, V. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6603 (1981).
- Schally, A.V., Baba, Y., Nair, R.M.G. & Bennett, C.A. *J. biol. Chem.* **246**, 6647 (1971).
- Szabo, M., Chu, L. & Frohman, L. *Endocrinology* **111**, 1235 (1982).
- Frohman, L.A., Szabo, M., Berelowitz, M. & Stachura, M.E. *J. clin. Invest.* **65**, 43 (1980).

Learning and development

Activity-mediated neural change

from Jeff W. Lichtman and Dale Purves

MOST neurobiologists agree that neural activity causes long-lasting changes in the nervous system. The argument is as follows: experience modifies behavioural responses, therefore the electrical activity of neural elements must somehow lead to changes in the concatenations of neurones which mediate various behaviours. The question is: how?

The problem has been examined in two different and seemingly unrelated contexts. One approach has been to study synaptic efficacy in adult animals. Activity of synapses in several neural systems, including the hippocampus¹, the cerebellum² and some invertebrate ganglia³, modifies synaptic function for relatively long periods. For example, E. Kandel, J. Schwartz and their colleagues³ have described changes in the function of synapses between identified neurones in the abdominal ganglion of the mollusc *Aplysia* that occur in parallel with long-lasting changes in a simple behaviour (the gill-withdrawal reflex); they have recently proposed a detailed molecular model, based on synaptic modulation by cyclic AMP-dependent protein phosphorylation, to explain these effects.

A quite different context for exploring the ability of activity to induce long-lasting changes in the nervous system is development (see ref. 4). Here the focus has been on patterns of connectivity rather than on the function of individual synapses. Evidence that activity influences connections in the developing brain was first provided by D. Hubel and T. Wiesel⁵⁻⁸ who showed that interference with early visual experience has profound effects on at least two properties of the adult visual cortex. In newborn cats and monkeys the inputs from the two eyes gradually segregate, leaving bands which are innervated alternately by the right or the left eye. If one eye is covered by a patch in young animals, the columns it supplies are narrower than the columns served by the active eye⁸. The same phenomenon has been described in lower vertebrates where a columnar segregation can be induced in the optic tectum⁹⁻¹¹. In addition, visual deprivation can affect the proportion of neurones driven by one eye or the other⁵⁻⁷ (see below). These observations leave little doubt that many initial connections between central neurones and their inputs can be made or broken under the influence of activity.

Although some theoretical treatments have attempted to explain these changes in neural connectivity (for example, refs 12,13), until recently little experimental effort has been directed towards understanding the underlying mechanisms. The present resurgence of interest stems from

the discovery of similar changes in connectivity during normal development in the peripheral nervous system. Muscle fibres and some autonomic ganglion cells are now known to receive inputs from a greater number of axons at early developmental stages than in maturity^{14,15}. The elimination of some initial connections evidently results from a competitive withdrawal of axonal branches during a period when the number of synapses is actually increasing (that is, each axon elaborates more synaptic contacts on fewer target cells). A variety of experiments has shown that the course of this competitive rearrangement is influenced by activity: reducing the activity of the axons innervating muscle fibres by blocking action potentials or neuromuscular transmission slows the process of synapse elimination¹⁶⁻²⁰ whereas augmented activity increases the rate of rearrangement^{21,22}. Experiments using activity-blocking agents in the visual system have further documented the similarity of central and peripheral events. M. Stryker²³ has shown that if visual activity is abolished by repeated intraocular injections of tetrodotoxin (a strategy similar to blocking a motor nerve), the normal segregation of ocular dominance columns is delayed in much the same way that paralysis delays synapse elimination in the periphery. These findings not only argue for the generality of activity-related changes in connectivity, but also offer the possibility of detailed analysis in systems far simpler than the visual cortex.

One of the major issues that remains unsolved is the influence of timing on the competitive rearrangement of developing neuronal connections. Hubel and Wiesel found that an overall reduction of sensory activity produced by bilateral visual deprivation leads only to minor changes in the proportion of binocular neurones in adult animals (ref. 5, but see ref. 24). If, on the other hand, one eye or the other is alternately occluded (or the animal made cross-eyed so that the two eyes see different scenes), the number of binocular cells in the adult cortex is sharply reduced⁷. The reason for these effects may be interference with the normal temporal relation of activity among potentially competing axons. Many cells in the primary visual cortex are driven by inputs activated by stimulation of corresponding points on the two retinas. Since corresponding retinal points see the same part of the visual world, the inputs to these binocular neurones must be active at more or less the same time. G. Blasdel and J. Pettigrew²⁵ have shown that little competition occurs unless the activity arising from stimulation of each eye is separated by at least 10 seconds. Based on

the effects of alternating occlusion and squint, Hubel and Wiesel suggested⁷ that synchronous activity might abrogate competition between inputs that would otherwise interact.

For technical reasons, experiments in the periphery have so far involved only changes in the amount of activity, not in its relative timing. Thus, it may be that activity-related changes in the rate of synapse elimination in muscle and autonomic ganglia arise from an acceleration or a slowdown of maturation rather than as a direct effect of activity on synaptic competition. There are clues, however, that the temporal relationship of activity between competitors is also important in the generation of peripheral patterns of innervation. It is difficult to explain, for example, why an autonomic axon normally establishes many boutons on one ganglion cell rather than distributing the same number of synapses among a larger number of equally accessible target neurones. The capture of individual neurones by numerous synaptic contacts arising from the same axon requires that, during competition, these boutons be identified as members of the same set²⁶⁻²⁸. One element in the explanation of this phenomenon may be that the terminals of any one axon receive a trophic advantage conferred by their synchrony (see ref. 25).

Interestingly, the importance of appropriately timed activity on neural connections was suggested long ago by the psychologist D. Hebb²⁹. Whether by foresight or good luck, Hebb's 'principle' will probably turn out to be an important agent of connectivity; however, the idea seems as likely to find vindication in the context of

- Lynch, G. & Shubert, P. A. *Rev. Neurosci.* 3, 1 (1980).
- Andersen, P. *Trends Neurosci.* 5, 324 (1982).
- Kandel, E.R. & Schwartz, J.H. *Science* 218, 433 (1982).
- Harris, W.A. *A. Rev. Physiol.* 43, 689 (1981).
- Wiesel, T.N. & Hubel, D.H. *J. Neurophysiol.* 26, 1003 (1963).
- Wiesel, T.N. & Hubel, D.H. *J. Neurophysiol.* 28, 1029 (1965).
- Hubel, D.H. & Wiesel, T.N. *J. Neurophysiol.* 28, 1041 (1965).
- Hubel, D.H., Wiesel, T.N. & LeVay, S. *Phil. Trans. R. Soc. B278*, 377 (1977).
- Law, M.I. & Constantine-Paton, M. *J. Neurosci.* 1, 741 (1981).
- Fawcett, J.W. & Willshaw, D.J. *Nature* 296, 350 (1982).
- Meyer, R.L. *Science* 218, 589 (1982).
- Stent, G.S. *Proc. natn. Acad. Sci. U.S.A.* 70, 997 (1973).
- Swindale, N.V. *Proc. R. Soc. B208*, 243 (1980).
- Purves, D. & Lichtman, J.W. *Science* 210, 153 (1980).
- Brown, M.C., Holland, R.L. & Hopkins, W.G. *Br. Soc. dev. Biol. Symp.* 5, 245 (1981).
- Benoit, P. & Changeux, J.-P. *Brain Res.* 149, 89 (1978).
- Srihari, T. & Vrbova, G. *J. Neurocytol.* 7, 529 (1978).
- Thompson, W., Kuffler, D.P. & Jansen, J.K.S. *Neuroscience* 4, 271 (1979).
- Riley, D.A. *Brain Res.* 143, 162 (1978).
- Jackson, P.C. *Soc. Neurosci. Abstr.* 8, 709 (1982).
- O'Brien, R.A.D., Östberg, A.J.C. & Vrbova, G. *J. Physiol., Lond.* 282, 571 (1978).
- Thompson, W. *Soc. Neurosci. Abstr.* 7, 840 (1981).
- Stryker, M.P. *Soc. Neurosci. Abstr.* 7, 842 (1981).
- Swindale, N.V. *Nature* 290, 332 (1981).
- Blasdel, G.G. & Pettigrew, J.D. *J. Neurophysiol.* 42, 1692 (1979).
- Lichtman, J.W. & Purves, D. *Br. Soc. dev. Biol. Symp.* 5, 233 (1981).
- Lichtman, J.W. *J. Physiol. Lond.* 302, 121 (1980).
- Hume, R.I. & Purves, D. *Soc. Neurosci. Abstr.* 8, 709 (1982).
- Hebb, D.O. *The Organization of Behaviour* (Wiley, New York, 1949).

development as in the context of learning. This raises the fundamental question of whether activity-dependent mechanisms of learning and development are related. After all, the repertoire of behaviours based on patterns of connections laid down during normal development provide 'memories' no less important than memories acquired in postnatal life. What-

ever the ultimate answer to this question, advances in developmental neurobiology are eroding the traditional distinction between the establishment of neural connections and their subsequent modification. □

Jeff W. Lichtman and Dale Purves are in the Department of Physiology and Biophysics, Washington University School of Medicine, St Louis, Missouri 63110.

Exotic particles debate

Demon nuclei — a closer look

from Sverker Fredriksson and Magnus Jändel

FRANK CLOSE has recently in *News and Views* (*Nature* 296, 305; 1982) questioned our idea that a new form of matter has revealed itself in certain high-energy experiments (*Phys. Rev. Lett.* 48, 14; 1982). Or rather, Close quotes, uncritically, a paper by H. Lipkin (*Weizmann Inst. Rep.* WIS 82/2) which claims that the exotic six-quark particle ('demon deuteron') suggested by us cannot exist, since its properties would violate basic physical laws.

The demon deuteron consists of three diquarks (tightly bound pairs of quarks), not two nucleons ('triquarks') like the normal deuterium nucleus. The gross features of the demon are indeed strange, even for the world of elementary particles. The conjecture that it be almost stable leads, for example, to the requirement that the three diquarks have a net motion (oscillation) relative to the strong force field that makes the whole demon stay together. Lipkin claimed that 'such states cannot exist', and supported this statement with arguments from a non-relativistic potential model. The truth is, however, that such exotic modes of internal motion (called 'spurious' by Close) have been discussed for at least ten years in the literature on ordinary three-quark particles (baryons).

Realistic models for the structure of elementary particles must be founded on field theory, and such models require the existence of 'spurious' oscillations. A particularly illustrative example is given by the appealing 'soliton bag model' of T. Lee and others, where the confining field explicitly carries mass and momentum, which in turn make it oscillate relative to the constituent quarks. The interesting question is whether such states are experimentally identifiable among all the hundreds of other quark configurations.

In a revised version of the *Weizmann Institute Report* reviewed by Close, Lipkin has now (*Phys. Lett.* 117B, 437; 1982) changed his view, and claims that although demons might exist they would be experimentally uninteresting. This statement is based on the view that no particles with net oscillations of ordinary quarks have ever been observed, which can be understood if such excitations occur only at very high

energies, at which the systems are highly unstable. Consequently, a corresponding motion of three diquarks would be even less interesting. To support this view Lipkin quotes a particularly popular quark model, the MIT bag model, which is a convenient approximation of the more fundamental field theoretical approach. A closer look at that model shows, however, that all those exotic baryon excitations are not 'pushed up' to high masses, as stated by Lipkin, and some of them are indeed suggested in the literature to correspond to experimentally observed particles. One of those (the 'A 35') has its three quarks in a configuration that resembles our three-diquark demon, and also a mass below that

Frank Close replies

A three-body system has only two independent 'internal' orbital angular momenta. The original version of the demon deuteron required three p-wave orbitals. This is the essence of Lipkin's criticism: for a system comprised solely of three diquarks such a demon deuteron cannot exist.

These ideas, well known to nuclear physicists, have been successfully applied to quark systems during the past 15 years. Lipkin was one of the pioneers and has probably published more papers on quarks than any other theorist. The statement made by him in the circumstances of his original critique is, in my opinion, correct.

But quarks are confined — unlike nucleons — and this might provide a hidden further degree of freedom, such as in bag models. In such an eventuality, three genuine orbitals and a demon deuteron could result. That is the essence of the above article which now draws heavily on the bag model.

Their statement that such exotic modes "have been discussed for at least ten years . . ." gives a misleading picture. It is indeed true that ten years ago Drell and Johnson and, later, MIT bag theorists proposed such states, but it might be amusing to report that they thanked Lipkin for 'clarifying discussions' about these very spurious states. These models, *prima facie*, predict numerous such states which have never been confirmed in any experiment.

needed for the demon to be stable.

Except for such purely qualitative hints, there is not much to go on when trying to compute the demon mass, because current quark models seem incapable of estimating, with accuracy, the masses of new classes of elementary particles, that is, quark configurations other than those the models were designed to describe in the first run. In addition, the extra complication of a net quark oscillation is difficult to treat in any model, and the calculations quoted by Lipkin are subject to such additional assumptions that they are of little value for the demon case.

The debate about the comings and goings of demons has thus turned into a discussion of the ability of quark models to explain the normal elementary particles. We are, however, expecting the final judgement about demon nuclei to come from our experimental colleagues. The ongoing experiments at the Lawrence Berkeley Laboratory will hopefully settle the matter in around one year's time. Some interesting preliminary results have, in fact, been published recently by T. Liss *et al.* (*Phys. Rev. Lett.* 49, 775; 1982) and reported on in *New Scientist* (21 October 1982, p. 160). □

Sverker Fredriksson and Magnus Jändel are in the Department of Theoretical Physics, Royal Institute of Technology, S-100 44, Stockholm, Sweden.

This suggests that they do not exist or (which for practical purposes may amount to the same thing) are pushed to high masses where they are undetectable. This is the modern view. The Drell-Johnson model is now superseded in the opinion of most experts. There is even dispute as to whether such states exist in a bag model context: bag model applications to excited systems are still crude and in their infancy.

Finally, the statement that some such 'exotic' baryon excitations are experimentally observed. There is one baryon resonance observed that could be a collective excitation of quarks relative to the bag — if the bag is dynamical and not phenomenological, a point by no means understood — the 'A 35' mentioned above. However this baryon is perfectly consistent with being a conventional three-quark excitation (at the $N=3$ level in a harmonic oscillator potential), and to claim support for 'exotics' thereby is, I think, very risky.

These are the reasons why I quoted Lipkin (not uncritically as they claim) and why I am pessimistic that demon nuclei exist in the way that Fredriksson and Jändel now argue for them. If experimental selection panels have managed to find time and money to search for such entities, then we naturally all hope that some unexpected and clear signal will be seen. □

Frank Close is in the Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX.

Bone tools

The case of the clumsy cave-bears

from Paul G. Bahn

THE problems of deciding whether prehistoric bone assemblages were produced by men or by carnivores are currently very much to the fore in archaeology (see *Nature* 297, 607; 1982). They are having a special impact on our interpretation of the African Lower Palaeolithic^{1,2}, where the old notions of 'killer apes' and the 'osteodontokeratic culture' have been beating a retreat in recent work. Binford³ has even carried the attack to our interpretation of the European evidence by casting doubt on the bone tools found in the Middle Palaeolithic of Cueva Morin in Cantabrian Spain, although their validity is vigorously upheld by the excavator⁴.

The objects to be discussed here — three bone 'buttons' — appear at first glance to be man-made, but it is in fact highly probable, if not certain, that a carnivore was responsible. This conclusion should not, however, be taken as an indictment of all potential Middle Palaeolithic bone tools; it merely serves as a warning that the full context of a find should be investigated before claiming human manufacture.

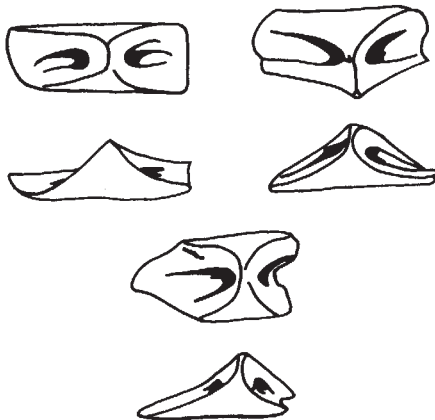
The 'buttons' were found in the late 1970s in the cave of Soulabé (Ariège, French Pyrenees) which has been excavated with exemplary skill for over 50 years by Léon Pales. It is only 10 metres from the famous cave of Malarnaud in which France's first fragment (a jaw) of Neanderthal man was discovered in 1888. Both caves open onto a cliff-face; Soulabé's upper gallery opens at 436 m altitude, and a second entrance lies 20 m below. The two galleries are linked by a vertical shaft⁵. Excavation has concentrated on the lower gallery and its prolongation. An important stratigraphy has been exposed, with 5 m of Middle Palaeolithic deposits. The two layers relevant here are the 'Ocre vii' a thick clay deposit assigned to the Würm II climatic phase⁶, where the first button was found in 1976; and the layer below, called 'Limon vert' thinner and assigned to the Würm I phase, where the other two buttons were found in 1977 and 1978.

The fact that only three 'buttons' have been found after so many years of excavation, that they were quite close to each other both horizontally and vertically and that they share a regularity and symmetry of form led to the notion that they had been carefully manufactured by Neanderthal man; indeed they have been described as well-made bone buttons similar to those of modern duffle-coats⁷.

Exactly similar objects have, however, long been known in the notorious and largely discredited 'Alpine Mousterian bone industry'^{8,9}. The claim¹⁰ that they were man-made was supported by many scholars, most notably Bächler⁹, who was

responsible for several myths of tools and bear-cults in the Alpine Mousterian¹².

The typical Alpine 'button' consists of a small portion of the shaft of a bear fibula with a double-bevelled 'bec de flûte' fracture at both ends and a bridge left in place in the centre. The three examples from Soulabé precisely fit this description (see the figure). A distribution map¹³ published in 1974 revealed that the buttons were known from only 13 caves, all of them in the Alpine/Circumalpine region and containing deposits of Mousterian age. As at Soulabé, the geographical and temporal concentration of the buttons, together with their regular and symmetrical shape, were taken as evidence of human manufacture. In addition, it was found possible to



Bone buttons.

fabricate such objects only on fresh or non-fossilized bone¹⁴.

F.-E. Koby, the implacable adversary of myths concerning cave-bears and Palaeolithic man, persistently claimed¹⁵ that the buttons, like other bone 'tools' of the Alpine Mousterian, were made accidentally, by bears blundering about in dark cave-depths and treading on the fragile bones of their antecedents. This opinion has found full support in a recent study by Giacobini⁹, whose detailed investigation of the morphology and biomechanical properties of bear fibulae showed that, depending on the condition of the bone and the position of the pressure-point, the double-bevelled fracture could very easily be produced by an animal stepping on the shaft. This explanation helps to account for the complete absence of incisions, traces of percussion or use-wear on the buttons.

Apart from man and bear, there is one further candidate for button-manufacturer — the hyena. Identical objects, in miniature, have been found among the bones of small mammals regurgitated by birds of prey¹⁶, and this has led to the theory that the normal buttons

may likewise have been made by a bone-crunching predator. The abundance of the buttons in some caves and their absence in others has therefore been linked to the presence or absence of hyenas⁸, although some see it as a cultural difference¹¹. Hence the rarity of buttons at Soulabé may be connected to the absence of hyena bones.

By and large the Alpine buttons have been found in 'bear-caves' — caves containing great quantities of cave-bear bones and sparse, if any, traces of human occupation. Indeed the buttons themselves, together with other alleged bone tools, are sometimes the only evidence for a human presence. Giacobini⁹ has postulated an inverse correlation between the abundance of buttons in a cave and the frequentation of the cave by man. Soulabé, like Malarnaud, was never more than a minor encampment for Middle Palaeolithic man, whereas the cave-bear used Soulabé intensively both as a 'nursery' and for hibernation⁵. Nevertheless the cave does contain a sparse lithic industry of mediocre quality, with tools of quartzite and flint, including some Levallois flakes; some teeth of Neanderthal man have also been found.

Unlike some Alpine caves, therefore, Soulabé was definitely used by man during the Mousterian. That fact, the relative proximity of its buttons to each other and the experiments of Schmidt may just make it feasible that these objects were man-made; however, the work of Koby and Giacobini, the observation of miniature buttons in pellets and the lack of definite human occupation in some Alpine caves make it highly probable that the objects were produced by bears blundering about in the 'nursery'. The Pyrenean finds make an important contribution to the debate in that they demonstrate that the buttons are by no means restricted to an imaginary bone-using Alpine Mousterian culture, but can occur in any cave much frequented by *Ursus spelaeus*. □

Paul G. Bahn is a Don Vicente Cañada Blanch Senior Research Fellow at the University of London.

- Behrensmeier, A. & Hill, A.P. (eds) *Fossils in the Making* (University of Chicago Press, 1980).
- Brain, C.K. *The Hunters or the Hunted?* (University of Chicago Press, 1981).
- Binford, L.R. *Bones: Ancient Men and Modern Myths* (Academic New York, 1981); *Curr. Anthropol.* 23, 177 (1982).
- Freeman, L.G. (ed.) *Views of the Past*, 29–51 (Mouton, The Hague, 1978); *Curr. Anthropol.* (in the press).
- Pales, L. *Annls Spéol.* 26, 195 (1971).
- Miskovsky, J.-C. et al. *Bull. Ass. Fr. Etud. Quat.*, 19 (1976).
- Clottes, J. *Gallia Préhist.* 22, 629 (1979).
- Jéquier, J.-P. *Le Moustérien Alpin, Révision Critique* (Eburudunum, Iverdon, 1976).
- Giacobini, G. *Preistoria Alpina* (in the press).
- Kellermann, D. *Abh. naturhist. Ges. Nürnberg* 20, 9 (1913).
- Bächler, E. *Das Alpine Paläolithikum der Schweiz* (Birkhäuser, Basel, 1940).
- Kurtén, B. *The Cave Bear Story* (Columbia University Press, New York, 1976).
- Fozzati, L. *Bull. Etud. préhist.* 6, 77 (1974).
- Schmidt, A. *Mannus* 26, 204 (1934).
- Koby, F.-E. *Verh. naturf. Ges. Basel* 54, 59 (1943); *Act. Soc. Jurassienne Emulation* 57, (1953).
- Mühlhofer, F. *Z. Wiener Prähist.* 24, 1 (1937); *Bayer. Vorgesch. Blätter* 24, 8.

Ecology

The structure of food webs

from Robert M. May

IN real communities of plants and animals, species are linked to each other — as predator and prey, competitors or mutualists — in a complex mesh of interactions. The schematic representation of this array of interconnections is called a food web. Although a good deal of scattered information about individual food webs has been available for some time, it is only in the last 10 years or so that people have begun a systematic attempt to understand what factors determine the structure of food webs. A recent workshop* on food webs, organized mainly by the ecologists of Oak Ridge and held amidst the fall foliage in the Smoky Mountains, was valuable both in drawing some strands of research together and in helping set an agenda for future work.

A primary resource for any theorizing about food webs is a careful catalogue of known webs whose structure has been elucidated. A relatively early such compilation of some 30 webs, published by Cohen¹ in 1978, stimulated a variety of theoretical analyses. Pimm², Rejmanek and Stary³, and others have extended and refined Cohen's catalogue; Briand's⁴ collection of 40 webs is probably the best available. At the workshop, Briand (University of Ottawa) reported that his catalogue now comprises 62 food webs and he doubts that many more will be found by searching in far corners of the existing literature. This 62-web collection breaks down into 19 from relatively constant environments versus 43 from fluctuating environments; 14 communities are tropical, 38 temperate and 10 are from high latitudes; 32 are from aquatic environments, 9 are terrestrial and 21 are from the land-sea interface.

As emphasized in the discussions, the compilation of such catalogues is beset with many difficulties, which are easier to state than to solve. First, for most purposes one wants a 'community' web, embracing all linkages among the co-occurring species; many published studies are of 'source' webs, which trace the connections upwards from a single resource species, or of 'sink' webs, which trace the links downwards from a single top predator. Second, there are problems of consistent definition of the entities in the web: often the studies articulate individual species of predators at the upper levels, but aggregate species (into 'copepods', or even more coarsely 'zooplankton') at lower levels. An extreme possibility is that different kinds of ecologists (for example, terrestrial people versus marine people) might systematically

follow different practices, thus creating the impression that terrestrial webs differed from marine webs, when in fact the differences were cultural ones among the investigators. Third, some studies include only those links between eaten and eater (the links along which energy or atoms flow up through the web), while others list links corresponding to direct competition or mutualism; which kind of web the theorist wants depends on the kind of analysis being undertaken. Fourth, and most important, at what point does one decide links are too weak or too unusual to list? Clearly, the flower-child ecologist's perception that to pluck a flower is to disturb a star is far too extreme, but there is a real dilemma in deciding, for example, whether Paine's⁵ rocky intertidal communities should embrace the owls that eat the mice that occasionally eat the odd limpet. Whether food webs should include all connections, or only the 'significant' ones, and how 'significant' is to be defined, stand out as a major unsolved problems⁵.

Several interesting patterns can be found from Briand's and earlier catalogues.

One is that the product of the number of species, S , and the connectance of the web, C (the ratio between the actual number of links and the maximum number that are topologically possible), is very roughly constant, at around 4–5, for all webs. Several people^{2,6} have observed that this is to be expected if, on average, all species interact directly with only a handful, $n \sim 4$ –5, of other species (then, for large S , $C \approx nS/S^2$ and $SC \approx n$). I, however, regard this rough constancy in SC as a surprising and interesting fact. Briand's further finding that SC tends to be systematically smaller for the 43 communities in fluctuating environments than for the 19 relatively constant ones is particularly pleasing to me, as it accords with one of the theoretical speculations that helped spark this work⁶.

Another pattern is that the number of omnivores — creatures consuming prey from two or more different trophic levels — in real food webs tends to be fewer than would be expected if the connections within the web were made randomly. This bears out a suggestion made by Pimm and Lawton^{2,7} on the basis of their studies of the dynamical stability of mathematical models for food webs. Discussion at the workshop made it clear, however, that this particular pattern, although statistically significant, is by no means universal.

While much theoretical activity seeks to find patterns common to all food webs, Briand used his 62-web collection to investigate differences among different environments. An intriguing finding is that food webs in two-dimensional environ-

ments (benthic; intertidal) tend to be wider and shorter — relatively more species of herbivores and fewer trophic levels — than those in three-dimensional environments (pelagic; three-dimensional terrestrial), where webs are relatively thin and long.

A more abstract pattern, first noted by Cohen¹, is that the feeding relationships depicted by 'consumer overlap graphs' can be represented in a one-dimensional niche space far more often than would be expected from chance alone. In the jargon of the graph theory trade, Cohen's observation is that food webs are usually 'interval graphs'; some intuitive appreciation of this notion of an interval graph is provided by the figure. Subsequent analyses of larger collections of food webs continue to confirm Cohen's finding that real webs are more commonly interval graphs than random assemblies would be.

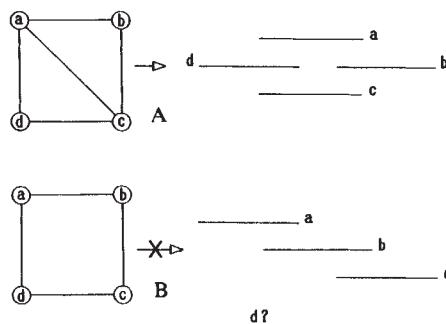
To my mind, the most exciting theoretical contribution at the workshop was by Sugihara (Oak Ridge National Laboratories), who has brought formal tools from graph theory and topology to bear on the analysis of food web structure. This work is not easily summarized. Crudely speaking, what Sugihara has shown is that the connected graphs representing relationships among resources are essentially free from 'holes'. To get some intuitive notion of what these holes are, consider a species that eats prey of small and intermediate sizes, and a second species that eats intermediate and large items; addition of a third species that eats small and large prey (but not intermediate ones) would create a 'hole' in the topological structure corresponding to the food web that contains these three species. Such holes are not found. In general, however, this 'unholiness' structural property of real food webs is more recondite than the above example suggests. Interestingly, Sugihara has shown that a hole-free graph necessarily possesses Cohen's interval property, provided also that the graph is not 'asteroidal' (roughly corresponding to the requirement that the graph is relatively compact, without too many waving tentacles). Thus intervality may be expected as a usual, but not invariable, corollary of 'unholiness'. Sugihara's work is interesting both as a nontrivial application of graph theory in biology, and as a phenomenological rule that begs for biological explanation⁸.

Side by side with our increasing knowledge about the empirical patterns exhibited by real food webs, there is a body of theoretical investigation of the properties of mathematical models of ecosystems. Most of the major papers have been published in *Nature*, so diligent readers of the journal will know that this body of work began about 10 years ago with studies of model webs in which the connections were made purely randomly^{6,9} (subject only to constraints set by the total

*A Food Web workshop was held at Fontana Village Inn, North Carolina on 25–27 October 1982. The proceedings will be available as an Oak Ridge National Laboratory Technical Report.

number of species S , the connectance C and the average interaction strength), and has become increasingly sophisticated as the effects of various kinds of constraints and rules for the assembly of model ecosystems have been explored^{2,7,10-15}. A variety of recent studies were reported at the workshop, some focusing on the successional patterns that emerge when webs are gradually built up by adding species according to some recipe (Glasser, University of Georgia; Post, ORNL; Drake, Purdue University), while others sought to deduce the biological rules that underlie the observed patterns in 'equilibrium' food webs (Pimm, University of Tennessee; Yodzis, Guelph University; King, ORNL).

Of particular interest is the light these theoretical studies shed on the factors determining the average length of food chains. In Briand's 62 food webs, the average chain length (the average number of links — eaten to eater — connecting top predators to basal resources) ranged from 5.9 to 1.9, with a mean of 2.9. The conventional explanation for the relative constancy and relative shortness of food chains is that at most 10 per cent or so of available energy is transferred from one trophic level to the next, and that this rapid attenuation forbids long chains in which predator is added upon predator. Pimm and Lawton¹⁶, however, have emphasized that such an explanation carries the implication that food chains should be significantly longer in highly productive systems (with a larger energy base) and in communities of cold- rather than warm-blooded species (because the efficiency of energy transfer between trophic levels is higher by one or two orders of magnitude for ectotherms than for endotherms); neither tendency appears to hold for real ecosystems. The workshop saw some debate on this point. Arguably, a tenfold increase in basal productivity could foster a tenfold increase in diversity, with different food chains culminating in predators so different that the expected top 'superpredator' is ecologically infeasible. Cohen (Rockefeller University) called for studies of the statistical distribution of food chain lengths (going beyond the present means and variances), and it would be nice to use such information, along with some objective index of the productivity of a given system, to see exactly what correlations do exist between distributions of food chain lengths and productivity. Similarly, a definitive comparison between the lengths of food chains in webs with ectotherms versus those with endotherms remains to be made. Pimm reported a good start on this task: beginning at the third trophic level (because things are confused at lower levels), he found the average length of food chains with invertebrate ectotherms to be 0.58 (for 12 webs), with vertebrate ectotherms to be 0.58 (again for 12 webs) and with endotherms 0.62 (for 8 webs). These facts seem to undercut an



The overlaps in resource use among four species (a, b, c, d) are depicted graphically on the right-hand side. The relationships described by consumer graph A can be reduced to the one-dimensional arrangement shown to the left: species a and c each overlap with the other three, while species b and d both overlap with a and c, but not with each other. Conversely, consumer graph B cannot be reduced to such a one-dimensional representation: as shown, b overlaps with a and c; d also overlaps with a and c; but a and c do not overlap with each other and therefore a one-dimensional representation is impossible. It follows that A is an 'interval graph' and B is not. (After ref. 8, following ref. 1.)

explanation of food web structure based on energetic considerations, as the energy transfer efficiency from level to level is around 20–50 per cent for invertebrate ectotherms, around 10 per cent for vertebrate ectotherms and typically less than 2 per cent for endotherms.

Pimm and others have offered the alternative explanation that dynamical considerations mould food web structure; in simulated webs, long food chains lead to severe population fluctuations that are inconsistent with long-term persistence. De Angelis *et al.*¹⁷ have observed, however, that such dynamical arguments are keyed to the characteristic doubling times of the slowest-changing populations, so that a tenfold speeding of metabolic rates throughout the ecosystem may — other things being equal — enable food chains to lengthen while the overall dynamical stability remains unchanged. The counter-argument is that top predators in most food webs are relatively large vertebrates, whose characteristic time scales for population change are all typically measured in years, leading very roughly to a common time scale for most webs. If true, this argument validates Pimm's approach, but leaves unanswered the deeper question of why top predators have this property. Perhaps it is because the behavioural constraints on top predators force them to be large, and thus long-lived, in which case we would have the remarkable situation that the behavioural ecology of individuals is intimately involved in explaining the gross structure of communities. Perhaps, on the other hand, different ecosystems do indeed have significantly different characteristic time scales. More analysis of the facts is needed.

Although the clash between mathemati-

cal models based on considerations of energy transfer (Yodzis and others) and those based on considerations of dynamical stability (Pimm and others) was a recurring theme of the meeting — with both classes of models being capable of a Procrustean fit to the facts at this relatively early stage — yet other classes of models remained in contention. Thus Hastings (Hofstra University) suggested that evolutionary pressures favour species consuming relatively low on the trophic ladder, where energy is, as it were, more abundant¹⁸; the counter-argument here is that even if less energy is on the hoof at higher levels, there is an evolutionary case for specializing on it if nobody else is competing for it. Nunney (UC Riverside), Abrams (University of Minnesota) and others observed that all the above theoretical work is based on the classic, but excessively simple Lotka–Volterra equations, and they argued that understanding of food web structure may necessarily involve such neglected details as time delays associated with complex life histories or with resource renewal, or the direct and indirect effects of nonlinearities in interactions between species¹⁹.

More broadly, a variety of other aspects of food web structure that are conventionally overlooked were identified at the workshop. Thus most studies omit all decomposers, or else treat them as a kind of consumer species; from some points of view, correct accounting for decomposer species converts an open system into a closed one, which can qualitatively alter some dynamical features. Parasites, including viruses, bacteria, protozoans and helminths, arguably regulate many natural populations; how should they be included in food webs? And how should we treat what Southwood has called 'tourist' species (species that are only accidentally or vagariously present, yet can on occasion consume a significant fraction of available resources)? Attempts to answer some of these questions lead naturally to a call to go beyond qualitative information about food webs — which species are linked to which — to quantitative information about the numerical magnitude of flows of energy or of nutrients (C, P, S, N and so on). One next step is the synoptic compilation of quantitative food webs, following the example set by Briand and others for qualitative ones.

It is over a century since Darwin made clear the basic way in which species originate. We are all aware of the continuing effort that goes into clarifying the details. But there remain several big questions: what determines the actual number of species, both locally and globally; what determines the relative abundance of species; how is the number of species of a given physical size determined; how are food webs structured? In view of their significance, I think these questions receive surprisingly little attention; to my knowledge, the meeting reported here was

the first devoted to food web structure. My suspicion is that, ultimately, these big questions are entwined, so that we may not be able truly to understand what determines the length of food chains without understanding what determines the number of species, their relative abundance and their spectrum of physical sizes. □

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, New Jersey 08544.

1. Cohen, J.E. *Food Webs and Niche Space* (Princeton University Press, 1978).
2. Pimm, S.L. *Food Webs* (Chapman & Hall, 1982).
3. Rejmanek, M. & Stary, P. *Nature* **280**, 311 (1979).
4. Briand, F. *Ecology* (in the press).
5. Paine, R.T. *J. Anim. Ecol.* **49**, 667 (1980).
6. May, R.M. *Nature* **238**, 413 (1972).
7. Pimm, S.L. & Lawton, J.H. *Nature* **275**, 542 (1978).
8. Sugihara, G. thesis, Princeton Univ. (1982).
9. Gardner, M.R. & Ashby, W.R. *Nature* **228**, 784 (1970).
10. Roberts, A.P. *Nature* **251**, 607 (1974).
11. Tregonning, K. & Roberts, A.P. *Nature* **281**, 563 (1979).
12. Yodzis, P. *Nature* **284**, 544 (1980).
13. Gilpin, M.E. & Case, T. *Nature* **261**, 40 (1976).
14. Yodzis, P. *J. theor. Biol.* **92**, 103 (1981).
15. De Angelis, D.L. *Ecology* **56**, 238 (1975).
16. Pimm, S.L. & Lawton, J.H. *Nature* **268**, 329 (1977).
17. De Angelis, D.L., et al. *Nature* **273**, 406 (1978).
18. Hastings, H.M. & Conrad, M. *Nature* **282**, 838 (1979).
19. Nunnery, L. *Am. Nat.* **115**, 639 (1980).

Atmospheric chemistry

Increasing atmospheric methane

from W.L. Chameides

In the past few years there have been several reports¹⁻⁴ that atmospheric methane levels, measured by the highly accurate gas chromatographic/flame ionization technique, have been increasing at the rate of 1.7 per cent per year since at least 1965 (refs 4,5). Now Craig and Chou⁴ have attempted to find out whether this is a short-term fluctuation or part of a long-term trend by taking polar ice cores and analysing air samples which were trapped in polar ice as far back as 27,000 BP. Their results lead to the startling conclusion that atmospheric methane concentrations remained fairly constant from 27,000 BP to AD 1580 at 0.7 p.p.m.v. (parts per million by volume) — about half the present value of 1.65 p.p.m.v. In 1580 methane levels began to rise, first at a rate of about 0.114 p.p.m.v. per century and then, from 1915, much more rapidly, at a rate of 2.5 p.p.m.v. per century, roughly equivalent to the rate of 1.7 per cent per year inferred from present-day measurements.

Should the approach of Craig and Chou prove to be valid — they make the assumption that methane trapped in ice is neither produced nor destroyed over tens of thousands of years, which, while not unreasonable, has yet to be validated — then their findings may have significant implications. An increase in methane from 0.7 p.p.m.v. to its present value of 1.66 p.p.m.v. is estimated to have caused an increase in global temperatures of about 0.23°C due to the atmospheric greenhouse effect⁶. The temperature rise is roughly 38

caused methane suddenly to increase towards the end of the sixteenth century after remaining constant for at least 25,000 years? Why did the rate of increase accelerate in this century?

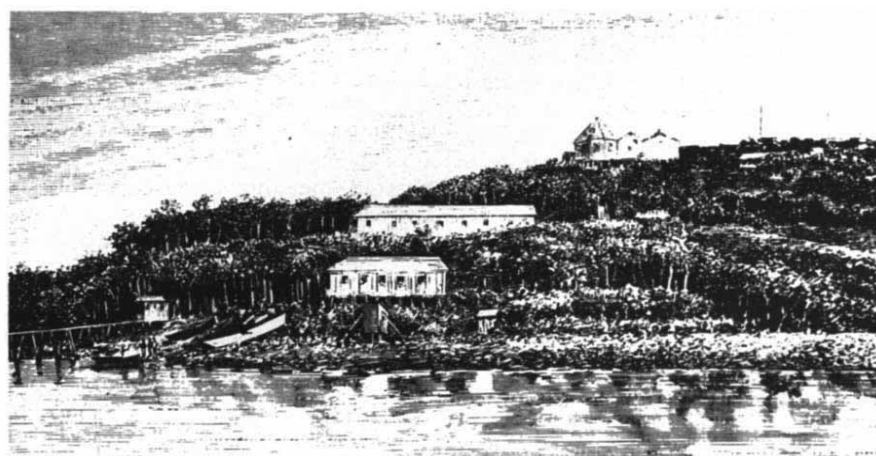
The primary sources of atmospheric methane are biogenic⁷ arising from the fermentation of organic material by anaerobic bacteria living in swamps, rice paddies, tropical rain forests and the intestinal tracts of livestock and termites. Methane is removed from the atmosphere photochemically by reaction with OH⁸ and approximately 200 million tons of methane are cycled through the atmosphere

per cent of the increase calculated to have occurred as a result of increases in atmospheric CO₂. Furthermore, since methane plays a central role in the chemistry of the atmosphere the increase may have caused changes in the levels of other atmospheric constituents such as OH, CO and O₃.

How can we explain the methane trends suggested by the data of Craig and Chou? What process or processes could have annually. Given a present-day methane abundance of 1.6 p.p.m.v., the atmospheric residence time for methane can be estimated at around 10 years.

One possible explanation of the increase is that atmospheric OH, which acts to remove CH₄ from the atmosphere, has decreased in concentration over this time period^{9,10}. Alternatively, the increasing CH₄ concentrations could be due to increasing biogenic production rates of CH₄. It is known, for instance, that as the world's population and demand for food has grown over the past centuries, the population of livestock and the total area of the world under rice paddy cultivation has also grown, implying a corresponding increase in methane production⁷. Should this latter explanation be correct it is likely that methane levels will continue their upward trend in the coming decades. We can only speculate as to the effects it will have on climate and global chemical cycles. □

W.L. Chameides is an Associate Professor of Geophysics in the School of Geophysical Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332.



100 years ago

THE FRENCH MISSION TO CAPE HORN

The members of the French Magnetic and Meteorological Expedition to Cape Horn have taken up their quarters at Crange Bay, and have already begun work. The accompanying illustration, reproduced from *La Nature*, after a photograph transmitted to the Paris Academy of Sciences, will give an idea of the aspect of the station occupied by the expedition. On the summit of the hill are the astronomical cabins, beside which are placed a pluviometer and an actinometer. The large house in the middle

distance forms the officers' quarters, while the lower building is for the sailors. Along the shore are other structures partly shown in the illustration, a stockade for the tidal register, and an isolated tent for absolute determinations.

The mission arrived at Orange Bay, Terra del Fuego, on September 6 last. They found the country marshy, and were compelled to select a wooden spot in order to obtain firm ground. No time was lost in erecting in-closings and installing the various instruments; and on September 26, the meteorological and magnetical observations were begun. From *Nature* **27**, 344, 8 February, 1883.

1. Rasmussen, R.A. & Khalil, M.A.K. *J. geophys. Res.* **86**, 9826 (1981).
2. Seiler, W. Personal communication (1982).
3. Blake, D.R. et al. *Geophys. Res. Lett.* **9**, 477 (1982).
4. Craig, H. & Chou, C.C. *Geophys. Res. Lett.* **9**, 1221 (1982).
5. Graedel, T.E. & McRae, J.E. *Geophys. Res. Lett.* **7**, 977 (1980).
6. Wang, W.C., Yung, Y.L., Laci, A.A., Tho, T. & Hansen, J.E. *Science* **194**, 685 (1976).
7. Enhalt, D.H. *Tellus* **26**, 58 (1974).
8. Levy, H. *Science* **173**, 141 (1971).
9. Sze, N.D. *Science* **195**, 673 (1977).
10. Chameides, W.L., Liu, S.C. & Cicerone, R.J. *J. geophys. Res.* **82**, 1795 (1977).

Calcium channel modulation by neurotransmitters, enzymes and drugs

Harald Reuter

Department of Pharmacology, University of Berne, 3010 Berne, Switzerland

Calcium channels in excitable membranes are of great importance for many cellular functions. Modulation of these channels by neurotransmitters and drugs regulates calcium influx into the cell and thereby alters the functional state of the cell. Recently it has become possible to measure properties of single calcium channels directly and to obtain evidence on mechanisms of their modulation.

AN increase in the free calcium ion concentration, $[Ca^{2+}]$, within stimulated cells underlies such important biological processes as muscle contraction^{1,2}, the secretion of transmitters and hormones^{3,4}, the regulation of enzyme activities⁵ and the control of membrane ion permeabilities⁶⁻⁸. Typically, the $[Ca^{2+}]$ of unstimulated cells is within the range 0.05–0.5 μM whereas that of the extracellular fluids is 1–10 mM. On electrical or chemical stimulation, the intracellular $[Ca^{2+}]$ increases by one to two orders of magnitude either as a result of Ca^{2+} influx across the cell membrane, or through the release of Ca^{2+} from intracellular stores, or both.

The purpose of this review is to summarize what has been learnt recently of the passive movement of Ca^{2+} ions through membranes, down the electrochemical gradient. Studies of this type have gained impetus with the advent of techniques enabling electrophysiological recordings to be made from intracellularly perfused single cells^{9,10} and from single ion channels in cell membranes¹¹. As most of our current knowledge of Ca^{2+} channels comes from the study of neurones and muscle cells, the review will be primarily concerned with these tissues. Ca^{2+} channels in the membranes of these cells are basically voltage dependent, although they can be modulated by neurotransmitters, hormones and drugs. Another class of Ca^{2+} channels, so-called 'receptor operated' channels, claimed to be present in a variety of other cells¹², will not be discussed because too little is known about their specific properties. The mechanisms by which the intracellular $[Ca^{2+}]$ is returned to that of the resting cell after stimulation will not be dealt with either. Interested readers are referred to articles in a recent book that deals separately with the mechanisms of extrusion of cell calcium by means of ATP-driven Ca-pumps¹³ or Na-Ca counter-transport¹⁴.

The first part of the review will summarize mechanisms involved in the regulation of membrane Ca^{2+} permeability, the second part will present the increasing body of evidence that shows modulation of Ca^{2+} channels by neurotransmitters, hormones and drugs, and the third part will deal with drugs and toxins that block Ca^{2+} channels.

Properties of Ca^{2+} channels

Passive movement of Ca^{2+} across plasma membranes is usually electrogenic, which means that, in favourable conditions, it can be measured as a membrane current (I_{Ca}). The driving force is the electrochemical gradient of the ion across the membrane, and the likely structures through which the ions permeate are 'pore'- or 'channel'-forming proteins embedded in the lipid bilayer of the membrane. Several extensive reviews have dealt with the apparent kinetic properties of Ca^{2+} currents in excitable membranes¹⁵⁻¹⁸, but only recently has it become possible to measure them with high precision and even to resolve the

current flowing through single Ca^{2+} channels. Biophysical studies of Ca^{2+} currents were previously hampered by structural complexities of the biological preparations and by the difficulty of separating I_{Ca} from other overlapping membrane current components. During the past few years these difficulties have been largely overcome by using single cells which can be internally perfused, by improved voltage-clamp techniques to measure membrane currents, and by better tools to eliminate current components other than I_{Ca} .

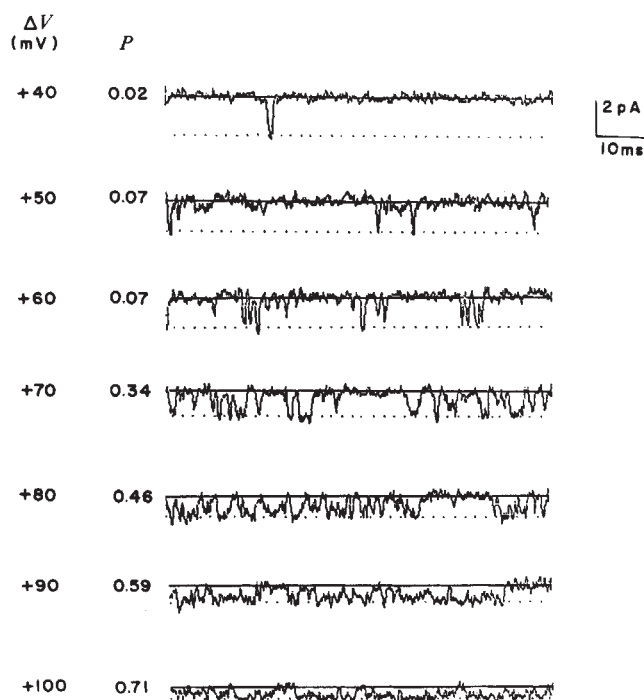


Fig. 1 Single Ca^{2+} -channel currents recorded from a cultured rat heart cell grown in a medium containing 0.1 mM 8-bromocyclic AMP. Voltage dependence of single channel currents, with Ba^{2+} (95 mM) as charge carriers. Membrane potential was stepped by the indicated step size (ΔV) from a holding potential of about -70 mV. The solid lines indicate the baseline after leak subtraction, the dotted lines indicate the measured open channel current. Beside each record is the average probability (P) that the channel was open during the voltage step. P was calculated from $I = NP_i$, where I is the mean current found by time-averaging over 100 ms the current flowing through the open Ca^{2+} channel, N is the number of functioning Ca^{2+} channels ($N = 1$ in Fig. 1) and i is the single channel current. i becomes smaller during depolarization because of the reduction in driving force (from ref. 35).

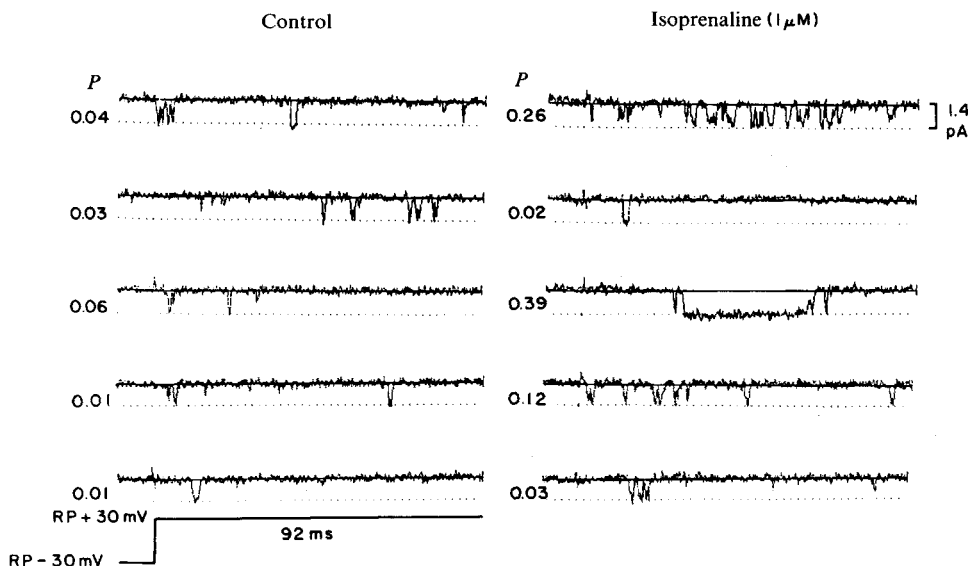


Fig. 2 Effect of isoprenaline ($1 \mu\text{M}$) on Ca^{2+} -channel currents (Ba^{2+} as charge carriers) measured from a cultured rat heart cell as in Fig. 1 (but without 8-bromo-cyclic AMP pretreatment). The holding potential was 30 mV negative to the resting potential (RP about -60 mV) and depolarizing clamp steps of 60 mV amplitude lasted 92 ms. The channel did not always open during 64 identical clamp steps. No openings (nulls) were observed in 21/64 traces in the control and in 5/64 identical traces in isoprenaline. Taking the nulls into account the average opening probability, P , of the channel increased significantly from 0.01 (controls) to 0.06 (isoprenaline), while the single channel current, i , remained the same (A. Cachelin, J. de Peyer and H.R., unpublished).

Ca^{2+} channels in excitable membranes are controlled by voltage-dependent gating, that is, their opening and closing kinetics are the result of changes in membrane potential. Gating is a complicated and incompletely understood mechanism¹⁹. Present evidence suggests that a 'voltage sensor' in the membrane, for example a protein group with dipole properties, which may be an integral part of an ion channel, reacts to the electric field. Any change in membrane potential will cause a reorientation of the charged sensor within the field and, therefore, a change in the ion flow through the channel. The movement of the gating particles will produce a measurable current¹⁹. Such 'gating currents' have also been described for Ca^{2+} channels in molluscan neurones^{20,21} and are believed to correspond to the activation of Ca^{2+} conductance (g_{Ca}) during depolarization.

The activation kinetics of g_{Ca} measured in single molluscan neurone somata seem to agree quite well with the kinetics of the gating currents²¹. However, there remains considerable doubt concerning the exact number of subunits and of charges displaced per opening of a single Ca^{2+} channel¹⁷. Furthermore, such a Hodgkin-Huxley-type kinetic model²¹ is not generally applicable to Ca^{2+} -channel activation and a different three-state sequential model, with two closed states and one open state, has been proposed²². Further experimentation will be necessary to resolve this problem before a more precise biophysical interpretation of the gating mechanism of Ca^{2+} channels can be

presented. It is an important problem not only for a better understanding of the molecular structure of the channel but also for the investigation of mechanisms of hormone and drug actions on Ca^{2+} channels some of which could be due to interference with channel gating.

There is a large electrochemical driving force for the inward movement of Ca^{2+} across cell membranes, because in unstimulated cells the equilibrium potential, V_{Ca} , according to the Nernst equation, is greater than $+120$ mV. However, the measured reversal potential, V_0 , where net current flow through Ca^{2+} channels is zero, is much more negative^{23,24}. In cardiac muscle^{23,25} and chromaffin cells²² K^+ and Cs^+ ions can move through Ca^{2+} channels in the outward direction and, in so doing, cause a shift of V_0 by 50–70 mV negative with respect to V_{Ca} .

Whether Ca^{2+} channels in neuronal membranes are also permeable to K^+ remains to be determined. However, it has been shown that Na^+ ions can move through Ca^{2+} channels, both in cardiac muscle and molluscan neurones, particularly with low extracellular $[\text{Ca}^{2+}]$ (refs 23, 26, 27). This has been explained by a binding site for divalent cations within the channel which controls its selectivity¹⁸. Despite the apparent incomplete selectivity of Ca^{2+} channels for Ca^{2+} , they are still among the most selective channels in excitable membranes. For example, the Na^+ channel is at least an order of magnitude less selective for Na^+ (ref. 28) than the Ca^{2+} channel is for Ca^{2+} (refs 17, 23).

Table 1 Comparison of steady-state properties of Ca^{2+} currents in various preparations

Tissue	Activation range (mV)	Inactivation range (mV)	Ca-dependent inactivation	Single-channel current (pA; -20 to -10 mV)	Refs
Neurones:					
Land snail	-50 to $+50$	-50 to $+50$	Yes	0.1; 0.2 (Ca^{2+} ; Ba^{2+})* 0.5; 0.8 (Ca^{2+} ; Ba^{2+})†	18, 32, 36, 37, 109 34, 38
<i>Aplysia</i>	-20 to $+30$	-40 to $+30$	Yes	?	30, 33, 110
<i>Limnaea</i>	-40 to $+30$	?	?	?	67
Squid giant synapse	-50 to $+30$	?	?	?	87
Dorsal root (rat)	-60 to -10	-120 to -10	?	?	69
Dorsal root (chick embryo)	-30 to $+20$	-30 to $+20$	Yes	?	78
Chromaffin cells	-40 to $+40$	—	Yes	0.1; 0.5 (Ca^{2+} ; Ba^{2+})* 0.9 (Ba^{2+})†	22
Muscles:					
Heart	-40 to $+10$	-50 to $+10$	Partially ?	0.8–1.4 (Ba^{2+})†	15, 23, 29, 35, 111, 112
Skeletal (insect)	-30 to $+20$	-40 to $+10$	Yes	?	31
Skeletal (frog)	-40 to $+10$	-50 to $+10$	?	?	39, 85

*Noise measurements. The ions in parentheses refer to the respective charge carriers.

†Direct single channel recording.

Conductance is usually turned on at more positive membrane potentials than Na^+ conductance. In cardiac muscle g_{Na} can be completely inactivated by membrane depolarization before any activation of g_{Ca} occurs^{15,29}. However, because outward currents in this tissue become smaller during depolarization (inward-going rectification), the relatively small Ca^{2+} current is sufficient to produce regenerative action potential activity. This feature is important in the heart for pacemaker activity and also for the genesis of arrhythmias²⁹.

Inactivation of Ca^{2+} channels is more complicated than that of Na^+ channels. In addition to voltage-dependent inactivation, as in Na^+ channels, there is, at least in some tissues, evidence for an 'inactivation' process that depends on the flow of I_{Ca} through the channels; the larger the current the faster the inactivation^{17,30,31}. In molluscan neurones Ca^{2+} ions have been shown to block Ca^{2+} channels from the inside at rather low concentrations^{18,32} which is probably the reason for I_{Ca} -dependent inactivation³³. These and other findings²² cast some doubt on the general applicability of the Hodgkin-Huxley formalism to Ca^{2+} channels.

Table 1 compares some features of Ca^{2+} channels in various tissues and includes data from the newly available technique¹¹

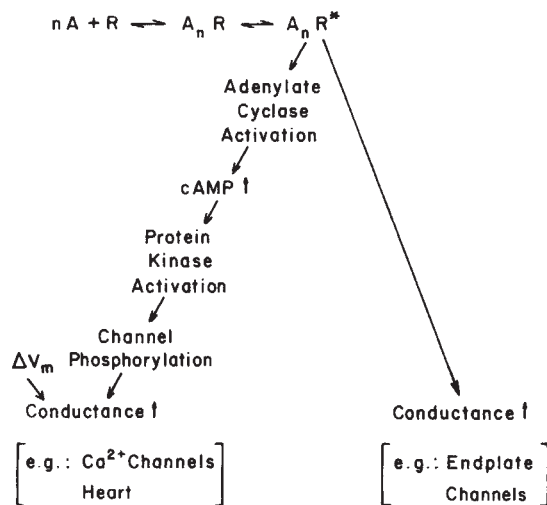
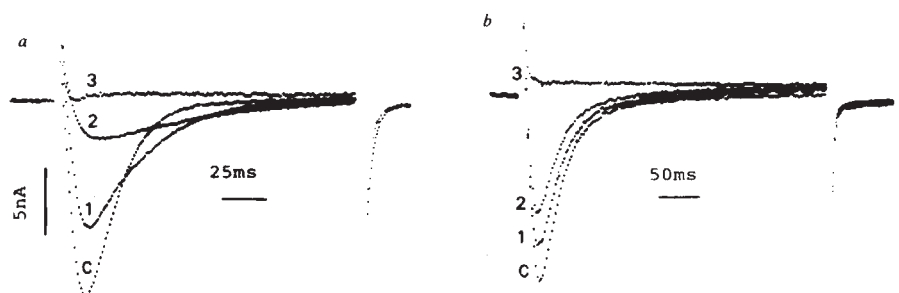


Fig. 3 Hypothesis for two different pathways for ion channel regulation in postsynaptic membranes. A, agonist; R, receptor; n , number of agonist molecules reacting with R; $A_n R^*$, activated agonist-receptor complex^{113,114}. The right-hand pathway applies to the action of a cholinergic nicotinic agonist, where it is known that the reaction of two acetylcholine molecules with the receptor subunit opens the channel¹¹³. The left-hand side may apply to modulation of Ca^{2+} channels by β -adrenergic agonists. After binding of the agonist to the β -receptor, the adenylate cyclase is activated via the guanine nucleotide-regulated N-unit¹¹⁴. This results in an increase in cyclic AMP which binds to the regulatory subunit of a protein kinase, thereby liberating the catalytic subunit of the enzyme⁵. As a result, phosphorylation of the ion channel, or of a protein closely associated with the channel, occurs. This causes a conformation of the channel allowing it to open on membrane depolarization, ΔV_m , with a higher probability than in the dephosphorylated state.

Fig. 4 Membrane Ca^{2+} currents recorded from differentiated N1E 115 neuroblastoma cells in voltage-clamp conditions. Depolarizing voltage steps of 70 mV amplitude and 170 ms duration were applied from a holding potential of -80 mV. *a*, Time course of the inhibitory effect of 2×10^{-4} M veratridine: C, control; traces 1, 2 and 3 were obtained 1, 2 and 5 min after the addition of the drug. *b*, Time course of the blocking effect of 2×10^{-7} M batrachotoxin: C, control; traces 1, 2 and 3 were measured 10, 15 and 30 min after addition of the toxin (from ref. 96).



that allows the measurement of the properties of single Ca^{2+} channels in membranes. Current flow through single channels can be recorded by means of the patch clamp method¹¹. Three recent reports on Ca^{2+} channels in *Helix* neurones³⁴, chromaffin²² and cardiac cells³⁵ show that single-channel currents (Table 1) and slope conductances are substantially larger than previously estimated from noise analysis^{36,37}. Figure 1 shows recordings of Ba^{2+} currents through a single Ca^{2+} channel in a cardiac cell³⁵. The probability (P) that the channel opens increases steeply with depolarization. Over the voltage range tested the slope conductance of the channel is approximately ohmic (usually 5–15 pS; 25 pS in Fig. 1). Typically, the channel opens in bursts and the burst lengths increase with voltage, while the average duration of channel openings (~ 1 ms) is only moderately voltage dependent^{22,35}. It seems that the burst kinetics rather than the mean open times of a channel determines the rate by which the total current, I_{Ca} , turns on at any given depolarization³⁵. Inactivation results from reduced probability that a channel is available to open and not from a reduction of single-channel current³⁵. From a comparative study on Ca^{2+} channels in widely different tissues it has been suggested that elementary Ca^{2+} currents are very similar³⁸, although possible differences seem to exist in the mechanisms in inactivation^{22,35}.

By knowing total membrane Ca^{2+} current density, single-channel current and membrane capacity, one can calculate a density of 2,000–10,000 functional Ca^{2+} channels per cultured rat cardiac cell or 0.1–0.5 channels per μm^2 membrane surface area (5–15 channels per μm^2 in chromaffin cells²²). In frog skeletal muscle, Ca^{2+} channels seem to be primarily located in the T-tubulus³⁹.

Modulation of Ca^{2+} channels

Cardiac muscle. Probably the first evidence that Ca^{2+} currents can be modulated by neurotransmitters and drugs was obtained in cardiac muscle (for reviews see refs 40, 41). Adrenaline and noradrenaline increase I_{Ca} in mammalian^{42–47} and frog^{48–50} myocardium. This causes an elevation of the plateau height of the cardiac action potential^{35–38,40–51}, an effect which can be mimicked by the injection of cyclic AMP into myocardial fibres^{52,53}, by the application of mono- or dibutyl-cyclic AMP or phosphodiesterase inhibitors^{43,50,54}, or by injecting single myocardial cells with the catalytic subunit of cyclic AMP-dependent protein kinase⁵⁵. The catecholamine effect on I_{Ca} and plateau height can be inhibited by β -adrenoreceptor antagonists⁴³.

Detailed kinetic analyses of I_{Ca} in the presence of β -adrenergic neurotransmitters have shown that neither the voltage-dependent kinetics of the current nor the reversal potential, V_0 , are much affected by the transmitters, but that the maximal conductance, g_{Ca} , is greatly increased^{44,46,47}. We suggested that this effect was explained by an increase in the number of Ca^{2+} channels available to open during depolarization⁴⁴. However, it could also be explained by an increased probability of individual channels to open at any given voltage⁴¹. Furthermore, in the earlier experiments the evidence in favour of unaltered single channel currents was indirect. Recent measurements of currents flowing through single Ca^{2+} channels have confirmed that isoprenaline does not affect the size of the elementary currents³⁵. Figure 2 illustrates this result and also

shows that, on the average, the probability, P , of the channel entering an open state is increased by isoprenaline. There are more frequent occurrences of bursts during identical step depolarizations to always the same potential, and possibly a slight lengthening of the mean open time of the channel³⁵. In other experiments with more than one channel in a patch, at a given voltage an increase in the average number of open channels was observed. Averaging many such single-channel current traces by means of a computer showed a substantial increase in the resulting mean current amplitude by isoprenaline without an obvious change in the turn-on kinetics (A. Cachelin, J. De Peyer and H.R., unpublished). The conclusion from these results is that β -adrenoreceptor agonists enhance I_{Ca} by increasing the probability of Ca^{2+} channels to open during depolarization.

The long-held belief^{43,44,51,54-56} that phosphorylation of Ca^{2+} channels, or of membrane proteins closely associated with the channels, is involved in the β -adrenergic effect on I_{Ca} in the heart (Fig. 3), presumably by increasing the probability of the channels entering an open state, has gained support by a recent report⁵⁷ showing cyclic AMP-dependent phosphorylation of a specific membrane protein and a concomitant stimulation of Ca^{2+} flux in sarcolemmal vesicles. Furthermore, histamine, through H_2 -receptors, not only stimulates cardiac adenylate cyclase⁵⁸, but also enhances Ca^{2+} current^{59,60}, possibly by the same sequence of events as illustrated in Fig. 3. However, it should be remembered that in addition to cyclic AMP-dependent protein kinase there are several other protein kinases in the cell (for example, Ca^{2+} -calmodulin dependent protein kinase, protein kinase-C requiring diacylglycerol, phosphatidylserine and Ca^{2+} ; see ref. 5 for a review), all of which could in principle phosphorylate ion channels in membranes. It will be interesting to see whether any of these enzymes is involved in the increase of I_{Ca} in the heart by cardiac glycosides^{61,62}, because a difference from the enhancement of I_{Ca} by adrenaline is suggested by the fact that only the glycoside effect is suppressed by the injection of EGTA into the cell⁶². Furthermore, cardiac glycosides do not stimulate adenylate cyclase⁶³.

The ability of muscarinic cholinergic agonists (acetylcholine, carbachol) to inhibit I_{Ca} in the heart^{64,65} has not been analysed in detail, although they seem to reduce $\bar{g}_{Ca}$ (ref. 65). Single-channel analysis will show whether or not this can be attributed to a decreased opening probability of the channels, an effect which would be exactly opposite to that seen with adrenergic agonists. It is of interest that the cholinergic inhibitory effect on I_{Ca} is particularly prominent when muscarinic drugs depress cyclic AMP formation after a preceding adrenergic stimulation⁶⁶. Whether an increase in cyclic GMP is involved in the reduction of I_{Ca} by muscarinic agonist is another open question⁵³.

Neurones. The experience of various investigators indicates that Ca^{2+} currents in internally perfused neurones are rather labile^{18,67,69}. They decrease continuously over a period of several minutes. In dorsal root ganglion neurones of newborn rats the decay can be prevented or even reversed by the addition of cyclic AMP, ATP and Mg^{2+} to the internal dialysis medium^{18,69}. In neurones from suboesophageal ganglia of *Helix aspersa*, intracellular perfusion with cyclic AMP, or extracellular application of ATP or its non-hydrolysable congener AMP-PNP, at nanomolar concentrations caused a slight but reproducible increase in I_{Ca} (ref. 68). These results have been interpreted^{18,69} in terms of the phosphorylation hypothesis of Ca^{2+} channels as described for cardiac muscle. However, at least in other molluscan neurones the effects are not easily reproducible⁶⁷.

An interesting and clearer mechanism whereby I_{Ca} is indirectly increased has been detected in sensory neurones of *Aplysia californica*. Initially it was shown that serotonin which causes an increase in cyclic AMP in these cells, or intracellular injection of cyclic AMP produce an enhancement of action potential amplitude and duration⁷⁰. Although, in analogy to cardiac muscle, an increase in I_{Ca} was suspected to be respon-

sible for these effects⁷⁰, it was later shown that I_{Ca} was little changed, whereas an overlapping K^+ conductance was reduced^{71,72}. This effect was explored in detail by using the patch clamp method to analyse the modulation of individual K^+ channels. Extracellular application of serotonin to the cell body or intracellular injection of cyclic AMP caused a prolonged and complete closure of individual K^+ channels⁷³. This accounts for the larger size and longer duration of the action potential and hence a prolongation of Ca^{2+} influx into the cell. As Ca^{2+} influx is intimately related to neurotransmitter release^{3,4}, this provides an explanation for the facilitated synaptic transmission observed in the same conditions⁷⁰⁻⁷². Facilitation of transmitter release by adrenaline and dibutyryl-cyclic AMP from bullfrog sympathetic ganglia has also been reported⁷⁴, but whether the mechanism is the same as that in *Aplysia* sensory neurones remains to be seen.

Much less is known about the mechanism by which α -adrenoreceptor stimulation in vertebrate neurones produces inhibition of I_{Ca} , although it is possible that a reduction of cyclic AMP is involved, because α_2 -adrenoreceptor agonists inhibit adenylate cyclase activity⁷⁵. In sympathetic neurones of the rat⁷⁶ and in cultured chick sensory neurones⁷⁷, noradrenaline inhibits Ca^{2+} -dependent action potentials. In both types of neurones this inhibition is due to a reduction of I_{Ca} by the neurotransmitter. The depression of I_{Ca} is the result of a reduction of the Ca^{2+} conductance, $\bar{g}_{Ca}$, while the kinetics of the current are unchanged^{78,79}, an effect exactly opposite to that seen with β -adrenergic catecholamines in the heart⁴⁴. The inhibition seems to be an α -adrenergic effect because it could be antagonized by α -adrenoreceptor blockers, but not by β -blockers⁷⁶. However, no clear distinction has been made between α_1 - and α_2 -adrenoreceptors in these studies. In view of the prominent role of α_2 -adrenoreceptors in presynaptic inhibition of noradrenaline release^{80,81}, it would be interesting to know whether noradrenaline inhibits its own release by decreasing I_{Ca} in presynaptic nerve terminals. A clear correlation between presynaptic inhibition and a decrease of I_{Ca} , though by a different neurotransmitter, has been observed in *Aplysia* L₁₀ neurones⁸².

Other neurotransmitters that inhibit Ca^{2+} current in embryonic chick dorsal root ganglion cells grown in tissue culture include γ -aminobutyric acid, serotonin, somatostatin and enkephalin^{77,78}.

Functional consequences of Ca^{2+} channel modulation

How important are Ca^{2+} -channel modulation and/or variability of Ca^{2+} -channel densities for cellular functions? The kind of short-term modulation of Ca^{2+} channels described above is of considerable functional significance for a variety of processes that depend on fluctuations of the intracellular $[Ca^{2+}]$ during excitation. The size of these fluctuations depends in part on Ca^{2+} influx across the membrane. Because, however, release of Ca^{2+} from intracellular stores can also contribute to the fluctuations in the cytoplasmic $[Ca^{2+}]$, the direct impact of variation in Ca^{2+} influx cannot always be assessed. An example is given by comparing the significance of Ca^{2+} influx through Ca^{2+} channels in frog heart and frog skeletal muscle. In frog heart contractile force seems to be directly related to Ca^{2+} influx, and increase in force induced by β -adrenergic neurotransmitters can therefore be considered as a direct consequence of the alteration of Ca^{2+} influx by Ca^{2+} -channel modulation^{50,51}. By contrast, contractile force of skeletal muscle is independent of Ca^{2+} influx and entirely dependent on Ca^{2+} release from the sarcoplasmic reticulum^{83,84}. In this case Ca^{2+} influx through Ca^{2+} channels located in the T-tubulus may serve to slowly replenish these stores^{39,85}. Contractility of mammalian cardiac muscle seems to be regulated by both Ca^{2+} influx and Ca^{2+} release from the sarcoplasmic reticulum, and the respective contribution of each may be variable.

In neuronal tissues modulation of Ca^{2+} channels could be an important factor in the regulation of neurotransmitter release

from presynaptic nerve terminals. Several examples have already been mentioned. However, it should be stressed that the precise sequence of events following Ca^{2+} influx and finally leading to transmitter release by exocytosis is still unclear. Another consequence of Ca^{2+} influx is the activation of other ion channels by intracellular Ca^{2+} (refs 6–8). Again, the exact relationship between Ca^{2+} influx and changes in $[\text{Ca}^{2+}]$ at the inner membrane surface remains to be determined.

However, there is not only modulation but also uneven distribution of Ca^{2+} channels in neuronal membranes. For example, in squid giant axons the density of Ca^{2+} channels is unmeasurably small⁸⁶, while in the squid giant synapse large Ca^{2+} currents have been recorded⁸⁷. The nerve terminal is exactly the site where a high density of Ca^{2+} channels is required for transmitter release, while for action potential propagation along the axon these channels are of little significance. Another interesting aspect is the fact that in developing neuronal and muscle tissue Ca^{2+} -channel densities do not seem to be constant. At early stages of development the Ca^{2+} -channel densities are higher than at later stages⁸⁸. This kind of long-term regulation of channel densities, though poorly understood, may point to some fundamental involvement of Ca^{2+} influx in cell development.

Blockade of Ca^{2+} channels

Unfortunately there is no such specific inhibitor for Ca^{2+} channels as tetrodotoxin is for Na^{+} channels. However, there are several inorganic ions and organic compounds that block Ca^{2+} channels preferentially to other ion channels in excitable membranes. The inorganic ions that block the channels are Mg^{2+} , Mn^{2+} , Ni^{2+} , Co^{2+} , Cd^{2+} and La^{3+} . Because of the clinical relevance of organic 'Ca antagonists' their number is rapidly growing⁸⁹. So far, the most important are verapamil and its methoxy-derivative D600, nifedipine and diltiazem. Only verapamil and related substances have been analysed in detail in terms of their direct effects on I_{Ca} in cardiac muscle^{90–93}. One of the most interesting aspects of the blocking effects of this class of drugs is their voltage-dependent mode of action. As with the action of local anaesthetics on I_{Na} (refs 94, 95), the potency of these drugs on I_{Ca} is increased during depolarization of the membrane as well as during repetitive stimulation ('use-dependence')^{90–93}. It is not clear whether the association rates of the drug molecules with Ca^{2+} channels are facilitated if the channels open during depolarization, or whether the dissociation rates are reduced if the channels inactivate. There is no clear indication whether or not the blocking actions of nifedipine and diltiazem are also voltage dependent.

The voltage- and use-dependence of the blockade of Ca^{2+} channels by verapamil and its derivatives may be of clinical significance because the potency of these drugs in depolarized

cardiac tissue, for example near infarcted areas, should be greater than in normal tissue. Arrhythmias originating from such areas should be suppressed more effectively than excitation in normally polarized tissue.

The potency of these drugs to block Ca^{2+} channels in neurones seems to be considerably less than in cardiac cells^{17,96}, while the opposite is true for the blockade of Na^{+} channels by tetrodotoxin⁹⁷. Binding studies with radioligands may help to clarify this discrepancy⁹⁸. However, it should be noted that Ca^{2+} channels, in contrast to Na^{+} channels, are not functional in isolated membrane vesicles (ref. 22 and A. Cachelin, J. de Peyer and H.R., unpublished). Therefore, such binding studies should preferentially be done in intact cells.

Certain lipid-soluble toxins (veratridine, grayanotoxin, batrachotoxin) which were thought to be specific modulators of the Na^{+} conductance⁹⁹ have recently been shown to be potent inhibitors of the Ca^{2+} current in neuroblastoma cells (Fig. 4)⁹⁶. In particular, batrachotoxin blocks Ca^{2+} channels at concentrations that are at least an order of magnitude lower than those required for their interaction with Na^{+} channels⁹⁶. Thus, batrachotoxin seems to be the first natural toxin with considerably higher specificity for Ca^{2+} channels than for Na^{+} channels.

Conclusions

In view of the importance of Ca^{2+} ions in the control of cellular functions, it will be of considerable interest to explore in detail the mechanisms by which Ca^{2+} fluxes through cell membranes are regulated. Voltage-dependent gating of Ca^{2+} channels and its modulation may be only one option. Whether or not phosphorylation is a general pathway of Ca^{2+} -channel modulation and whether, for example, opioids and related neuropeptides act by such a mechanism remains to be seen^{100,101}. Ca^{2+} permeability changes that are independent of voltage and are linked to phospholipid metabolism may turn out to be involved in transmitter and hormone actions different from those reviewed here^{12,102}. A voltage-dependent Na–Ca exchange seems to be yet another route of controlling intracellular Ca^{2+} (refs 14, 103). Specific inhibitors for the different Ca^{2+} pathways in cell membranes would not only help to sort out the relative functional importance of each system, but could also have considerable clinical impact. Calmodulin as an intracellular Ca^{2+} receptor and its importance in mediating certain Ca^{2+} -dependent cellular processes have been reviewed before^{5,104} and are beyond the scope of this article.

In general, the modulation of ion channels in membranes seems to be a subject of increasing interest^{105–107}, which is not surprising in view of their important role in the transmission of electrical and chemical signals and their possible involvement in complex events such as cell development⁸⁸ or behaviour¹⁰⁸.

This work was supported by the Swiss NSF.

- Ebashi, S. & Endo, M. *Prog. Biophys. molec. Biol.* **18**, 123–128 (1968).
- Allen, D. G. & Kurihara, S. *Eur. Heart J.* **1**, Suppl. A, 5–15 (1980).
- Katz, B. *The Release of Neural Transmitter Substances* (Liverpool University Press, 1969).
- Baker, P. F. & Knight, D. E. *Phil. Trans. R. Soc. B* **296**, 83–103 (1981).
- Cohen, P. *Nature* **296**, 613–620 (1982).
- Meech, R. W. A. *Rev. Biophys. Bioengng* **7**, 1–18 (1978).
- Colquhoun, D., Neher, E., Reuter, H. & Stevens, C. F. *Nature* **294**, 752–754 (1981).
- Maruyama, Y. & Petersen, O. H. *Nature* **300**, 61–63 (1982).
- Lee, K. S., Akaïke, N. & Brown, A. M. *J. Neurosci. Meth.* **2**, 51–78 (1980).
- Kostyuk, P. G., Krishtal, O. A. & Pidoplichko, V. I. *J. Neurosci. Meth.* **4**, 201–210 (1981).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
- Berridge, M. J. *Handbk exp. Pharmac.* **58** Pt II, 227–270 (1982).
- Schatzmann, H. J. in *Membranes and Transport* Vol. 1 (ed. Martonosi, A.) 601–605 (Plenum, New York, 1982).
- Reuter, H. in *Membranes and Transport* Vol. 1 (ed. Martonosi, A.) 623–631 (Plenum, New York, 1982).
- Reuter, H. *Prog. Biophys. molec. Biol.* **26**, 1–43 (1973).
- Brown, A. M., Akaïke, N. & Lee, K. S. *Ann. N.Y. Acad. Sci.* **307**, 330–344 (1978).
- Hagiwara, S. & Byerly, L. A. *Rev. Neurosci.* **4**, 69–125 (1981).
- Kostyuk, P. G. *Biochim. biophys. Acta* **650**, 128–150 (1981).
- Almers, W. *Rev. Physiol. Biochem. Pharmac.* **82**, 96–190 (1978).
- Adams, D. J. & Gage, P. W. *Science* **192**, 783–784 (1976).
- Kostyuk, P. G., Krishtal, O. A. & Pidoplichko, V. I. *J. Physiol., Lond.* **310**, 403–421 (1981).
- Fenwick, E. M., Marty, A., & Neher, E. *J. Physiol., Lond.* **331**, 599–635 (1982).
- Reuter, H. & Scholz, H. *J. Physiol., Lond.* **264**, 17–47 (1977).
- Adams, D. J., Smith, S. J. & Thompson, S. H. A. *Rev. Neurosci.* **3**, 141–167 (1980).
- Lee, K. S. & Tsien, R. W. *Nature* **297**, 498–501 (1982).
- Garnier, D., Rougier, O., Gargouil, M. & Coraboeuf, E. *Pflügers Arch. ges. Physiol.* **313**, 321–342 (1969).
- Kostyuk, P. G. & Krishtal, O. A. *J. Physiol., Lond.* **270**, 545–568 (1977).
- Hille, B. J. *J. gen. Physiol.* **58**, 599–619 (1971).
- Reuter, H. A. *Rev. Physiol.* **41**, 413–424 (1979).
- Tillotson, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1497–1500 (1979).
- Ashcroft, F. M. & Stanfield, P. R. *J. Physiol., Lond.* **330**, 349–372 (1982).
- Standen, N. B. *Nature* **293**, 158–159 (1981).
- Eckert, R. & Tillotson, D. L. *J. Physiol., Lond.* **314**, 265–280 (1981).
- Lux, H. D. & Nagy, K. *Pflügers Arch. ges. Physiol.* **391**, 252–254 (1981).
- Reuter, H., Stevens, C. F., Tsien, R. W. & Yellen, G. *Nature* **297**, 501–504 (1982).
- Akaïke, N., Fishman, H. M., Lee, K. S., Moore, L. E. & Brown, A. M. *Nature* **274**, 379–382 (1978).
- Krishtal, O. A., Pidoplichko, V. I. & Shakhvalov, Y. A. *J. Physiol., Lond.* **310**, 423–434 (1981).
- Brown, A. M., Camerer, H., Kunze, D. L. & Lux, H. D. *Nature* **299**, 156–158 (1982).
- Almers, W., Fink, R. & Palade, P. T. *J. Physiol., Lond.* **312**, 177–207 (1981).
- Reuter, H. in *The Slow Inward Current and Cardiac Arrhythmias* (eds Zipes, D. P., Bailey, J. C. & Elharrar, V.) 205–219 (Martinus Nijhoff, The Hague, 1980).
- Tsien, R. W. & Siegelbaum, S. in *Physiology of Membrane Disorders* (eds Andreoli, T. E., Hoffman, J. F. & Fanestil, D. D.) 517–538 (Plenum, New York, 1978).
- Reuter, H. *J. Physiol., Lond.* **192**, 479–492 (1967).
- Reuter, H. *J. Physiol., Lond.* **242**, 429–451 (1974).
- Reuter, H. & Scholz, H. *J. Physiol., Lond.* **264**, 49–62 (1977).
- Pappano, A. J. & Carmeliet, E. E. *Pflügers Arch. ges. Physiol.* **382**, 17–26 (1979).
- Kass, R. A. & Wiegand, S. E. *J. Physiol., Lond.* **322**, 541–558 (1982).
- Noma, A., Kotake, H. & Irisawa, H. *Pflügers Arch. ges. Physiol.* **388**, 1–9 (1980).
- Vassort, G. et al. *Pflügers Arch. ges. Physiol.* **309**, 70–81 (1969).
- Brown, H. F., McNaughton, P. A., Noble, D. & Noble, S. J. *Phil. Trans. R. Soc. B* **270**, 527–537 (1975).

50. Morad, M., Sander, C. & Weiss, J. *J. Physiol., Lond.* **311**, 585–604 (1981).
51. Niedergerke, R. & Page, S. *Proc. R. Soc. B* **197**, 333–362 (1977).
52. Tsien, R. W. *Nature new Biol.* **245**, 120–122 (1973).
53. Trautwein, W., Taniguchi, J. & Noma, A. *Pflügers Arch. ges. Physiol.* **392**, 307–314 (1982).
54. Tsien, R. W., Giles, W. & Greengard, P. *Nature new Biol.* **240**, 181–183 (1972).
55. Osterrieder, W. *et al. Nature* **298**, 576–578 (1982).
56. Tsien, R. W. *Adv. Cyclic Nucleotide Res.* **8**, 363–420 (1977).
57. Rinaldi, M. L., Le Peuch, C. J. & Demaille, J. G. *FEBS Lett.* **129**, 277–281 (1981).
58. Johnson, C. L. & Mizoguchi, H. *J. Pharmac. exp. Ther.* **200**, 174–186 (1977).
59. Sanchez-Chapula, J. *Eur. J. Pharmac.* **74**, 253–256 (1981).
60. Eckel, L., Gristwood, R. W., Nawrath, H., Owen, D. A. A. & Satter, P. *J. Physiol., Lond.* **330**, 111–123 (1982).
61. Weingart, R., Kass, R. S. & Tsien, R. W. *Nature* **273**, 389–392 (1978).
62. Marban, E. & Tsien, R. W. *J. Physiol., Lond.* **329**, 589–614 (1982).
63. Watanabe, A. M. & Besch, H. R. *Circulation Res.* **35**, 316–324 (1974).
64. Giles, W. & Noble, S. J. *J. Physiol., Lond.* **261**, 103–123 (1976).
65. Hino, N. & Ochi, R. *J. Physiol., Lond.* **307**, 183–197 (1980).
66. Biegon, R. L., Epstein, P. M. & Pappano, A. J. *J. Pharmac. exp. Ther.* **215**, 348–356 (1980).
67. Byerly, L. & Hagiwara, S. *J. Physiol., Lond.* **322**, 503–528 (1982).
68. Yatani, A., Tsuda, Y., Akaiki, N. & Brown, A. M. *Nature* **296**, 169–171 (1982).
69. Kostyuk, P. G., Veselovsky, N. S. & Fedulova, S. A. *Neuroscience* **6**, 2431–2437 (1981).
70. Klein, M. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3512–3516 (1978).
71. Shapiro, E., Castellucci, V. F. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 629–633 (1980).
72. Klein, M., Camardo, J. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5713–5717 (1982).
73. Siegelbaum, S., Camardo, J. S. & Kandel, E. R. *Nature* **299**, 413–417 (1982).
74. Kuba, E., Kato, E., Kumamoto, E., Koketsu, K. & Hirai, K. *Nature* **291**, 654–656 (1981).
75. Jakobs, K. H. *Molec. cell. Endocr.* **16**, 147–156 (1979).
76. Horn, J. P. & McAfee, D. J. *J. Physiol., Lond.* **301**, 191–204 (1980).
77. Dunlap, K. & Fischbach, G. D. *Nature* **276**, 837–839 (1978).
78. Dunlap, K. & Fischbach, G. D. *J. Physiol., Lond.* **317**, 519–535 (1981).
79. Galvan, M. & Adams, P. R. *Brain Res.* **244**, 135–144 (1982).
80. Langer, S. Z. *Br. J. Pharmac.* **60**, 481–497 (1977).
81. Starke, K. A. *Rev. Pharmac. Tox.* **21**, 7–30 (1981).
82. Shapiro, E., Castellucci, V. F. & Kandel, E. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1185–1189 (1980).
83. Armstrong, C. M., Bezanilla, F. M. & Horowicz, P. *Biochim. biophys. Acta* **267**, 605–608 (1972).
84. Lüttgau, H. C. & Spiecker, W. *J. Physiol., Lond.* **296**, 411–429 (1979).
85. Sanchez, J. A. & Stefani, E. *J. Physiol., Lond.* **283**, 197–209 (1978).
86. Meves, H. & Vogel, W. *J. Physiol., Lond.* **235**, 225–265 (1973).
87. Llinas, R., Steinberg, I. Z. & Walton, K. *Biophys. J.* **33**, 289–321 (1981).
88. Spitzer, N. C. A. *Rev. Neurosci.* **2**, 363–397 (1979).
89. Fleckenstein, A. A. *Rev. Pharmac. Tox.* **17**, 149–166 (1977).
90. Ehara, T. & Kaufmann, R. *J. Pharmac. exp. Ther.* **207**, 49–55 (1978).
91. McDonald, T. F., Pelzer, D. & Trautwein, W. *Pflügers Arch. ges. Physiol.* **385**, 175–179 (1980).
92. Trautwein, W., Pelzer, D., McDonald, T. F. & Osterrieder, W. *Naunyn-Schmiedeberg's Archs Pharmac.* **317**, 228–232 (1981).
93. Osterrieder, W., Pelzer, D., Yang, Q.-F. & Trautwein, W. *Naunyn-Schmiedeberg's Archs Pharmac.* **317**, 233–237 (1981).
94. Hille, B. *J. gen. Physiol.* **69**, 497–515 (1977).
95. Hondeghem, L. M. & Katzung, B. G. *Biochim. biophys. Acta* **472**, 373–398 (1977).
96. Romey, G. & Lazdunski, M. *Nature* **297**, 79–80 (1982).
97. Cohen, C. J., Bean, B. P., Colatski, T. J. & Tsien, R. W. *J. gen. Physiol.* **78**, 383–411 (1981).
98. Glossmann, H., Ferry, D. R., Lübbecke, F., Mewes, R., Hofmann, F. *Trends pharmac. Sci.* **3**, 431–437 (1982).
99. Honerjäger, P. *Rev. Physiol. Biochem. Pharmac.* **92**, 1–74 (1982).
100. Mudge, A. W., Leeman, S. & Fischbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 526–530 (1979).
101. Werz, M. A. & Macdonald, R. L. *Nature* **299**, 730–7333 (1982).
102. Mitchell, R. H. *Nature* **296**, 492–493 (1982).
103. Mullins, L. J. *Am. J. Physiol.* **236**, C103–C110 (1979).
104. Means, A. R. & Dedman, J. R. *Nature* **285**, 73–77 (1980).
105. Kehoe, J. S. & Marty, A. A. *Rev. Biophys. Bioengng* **9**, 437–465 (1980).
106. Hartzell, H. C. *Nature* **291**, 539–544 (1981).
107. Llinas, R. *Scient. Am.* **247**, 38–47 (1982).
108. Kandel, E. R. & Schwartz, J. H. *Science* **218**, 433–443 (1982).
109. Akaiki, N., Lee, K. S. & Brown, A. M. *J. gen. Physiol.* **71**, 509–531 (1978).
110. Adams, D. J. & Gage, P. W. *J. Physiol., Lond.* **289**, 143–161 (1979).
111. McDonald, T. F. A. *Rev. Physiol.* **44**, 425–434 (1982).
112. Tsien, R. W. & Marban, E. in *Advances in Pharmacology and Therapy* Vol. 3 (eds Yoshida, H., Hagihara, Y. & Ebashi, S.) 217–225 (Pergamon, Oxford, 1982).
113. Stevens, C. F. A. *Rev. Physiol.* **42**, 643–652 (1980).
114. Rodbell, M. *Nature* **284**, 17–22 (1980).

ARTICLES

From dwarfs to giants — signposts of galaxy formation

Joseph Silk*

Institut d'Astrophysique, 98 bis Boulevard Arago, 75014 Paris, France

A two-dimensional parametrization involving mean surface density and velocity dispersion suggests that many diverse properties of galaxies can be understood in terms of an evolutionary sequence operative during an early gas-rich protogalactic phase in which dissipation played a key role. Dwarf galaxies may be the fossilized link between primordial fluctuations and the giant galaxies observed today, relics of a past era of prolific galaxy formation.

To understand galactic evolution as well as we understand stellar evolution would be highly desirable. The complexity of galactic evolution with its many possible pathways makes this an elusive hope. Indeed, the physics of dissipative collapse of a massive gas cloud or ensemble of clouds is so complex that any direct attack using hydrodynamical studies is doomed at the outset. The amount of freedom in modelling supersonic turbulence and star formation is practically unlimited: when gravity is incorporated, any feasible problem degenerates to a parameter study, with little physical content. Although this has not deterred attempts at dissipative galaxy formation, the superficial successes of such models count for very little.

The final state of galactic evolution is observed. The time scale for secular dynamical evolution is currently very long, greatly exceeding a Hubble time for most isolated systems. Once evolution must have proceeded rapidly, however, when galaxies were predominantly gaseous. Even now, a merger may stimulate rapid evolution for a brief period. Consider a protogalactic system containing some specified mass in the form

of gas and stars. In effect, the stars are dynamical fossils; once some mass-fraction forms stars, dissipation effectively ceases for this component. Thus with care, we may hope to utilize the dynamical characteristics of the oldest components of galaxies, the spheroids, as a tracer of their prior dissipative history or, at least, to probe the final stages of the dissipative phase.

The initial state from which protogalactic evolution began is a matter of considerable conjecture, but theoretical work, nevertheless, leads to some rather specific alternatives. One may now have some cause for optimism, for if the initial state is reasonably well constrained and the final state is known, then, in principle, one can always compute evolutionary tracks. Moreover, knowledge of the detailed evolution may not be essential to an understanding of the relationship between, for example, different morphological types of galaxy.

The dynamical state of a stationary system is fully determined in the mean if we specify two dynamical parameters (three, if rotation is important). It is important that these be well observed parameters so that we can explore the implications for different types of galaxies. The parameters used here will be stellar velocity dispersion and mean surface density. The velocity dispersion is observed to vary hardly at all: spheroidal bulge

*Permanent address: Department of Astronomy, University of California, Berkeley, California 94720, USA.

components are isothermal, to a first approximation. The mean surface density will always be computed from the observed mean surface brightness, evaluated using integrated luminosity at the half-light radius, together with a value for the mass-luminosity ratio that is appropriate for the observed stellar component. The ratio of luminous mass (M_{lum}) to blue luminosity (L_B) includes contributions to the luminous mass within the Holmberg radius from H I and H₂. For Es and S0s, the value of M_{lum}/L_B is taken from Faber and Gallagher¹, whose solar neighbourhood value of this ratio is used for Sdms. For intermediate Hubble types, appropriate amounts of bulge are added, with $M_{\text{lum}}/L_B = 6$ corresponding to an old stellar population. The Hubble constant adopted is $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, where it enters these estimates.

It is well known that in a one-dimensional classification, involving either velocity dispersion σ or surface density Σ , there is separation of Hubble types, although with considerable overlap at a given value of the parameter. In the two-dimensional (Σ, σ) plane, the differences become more noticeable. Again, there is considerable dispersion at a given Hubble type, but mean lines through the observational correlations can be evaluated. To obtain (Σ, σ) relations, use is made of the luminosity-velocity dispersion or effective radius correlations, as described in Table 1. Measurements of the maximum rotational velocities of disks generally probe the dynamics of the spheroidal components. The core velocity dispersions are predominantly determined by the luminous stellar cores of ellipticals and S0s, so that the inferred surface densities refer to the luminous matter components. The velocity dispersions of galaxies in groups and clusters probe the dark matter distribution on megaparsec scales. The observed correlation between central galaxy density and galaxy velocity dispersion is used to derive a (Σ, σ) relation for groups and clusters of galaxies; again, the mean mass of the galaxy adopted is appropriate to the luminous mass content of these systems.

The various correlations are shown in Fig. 1, from which some immediate conclusions can be drawn. It should be emphasized that only mean correlation lines are plotted; in reality, there is very considerable overlap in areas occupied by different morphological types. Nevertheless, some galaxy types are separated along the Hubble sequence in order of increasing degrees of surface density and of bulge velocity dispersion from late through early-type spirals to SOs and ellipticals. Surface density and bulge velocity dispersion together conspire to yield a two-dimensional parametrization which hints at an underlying evolutionary scheme that may help unify the Hubble sequence.

Our choice of the (Σ, σ) parametrization was motivated by the following considerations of the theoretical aspects of galaxy formation. There are clear indications that dissipation had an important role in the formation of all galaxies, even including the ellipticals and dwarf spheroidals, as will be apparent from our analysis of Fig. 1. Consider a heterogeneous collection of gas clouds falling or coagulating together eventually to form a galaxy. Whether one favours the adiabatic pancake theory of galaxy formation², in which small-scale structure is suppressed and anisotropic collapse occurs to thin pancakes of mass $\sim 10^{14} M_{\odot}$ that subsequently fragment, or the hierarchical clustering theory, in which isothermal fluctuations of mass $\sim 10^6 M_{\odot}$ are the first to become Jeans unstable and collapse³, the basic building blocks for galaxies are of low mass. The minimum masses of the pancake fragments are $\sim 10^8 M_{\odot}$, and coagulate or accrete to form galaxies, whereas the $10^6 M_{\odot}$ precursor clouds cluster together hierarchically on larger and larger scales.

To explore the evolutionary significance of the (Σ, σ) diagram, we consider the two rival schemes of hierarchical clustering and pancake fragmentation for forming galaxies. As hierarchical clustering proceeds, larger and larger mass-scales condense out of the expanding universe. In the absence of dissipation, the mass scale that first becomes non-linear at epoch t increases as $t^{4(n+3)^{-1}}$, where n is the initial fluctuation spectral index (the amplitude of density fluctuations $\propto M^{-1/2-n/6}$). One

Table 1 Surface density-velocity dispersion correlations*

		Adopted M/L	Ref.
From observed surface brightness			
Luminous E	$10^3 (M/L) M_{\odot} \text{ pc}^{-2}$	8	17
Dwarf Irr	$10^{1.2} (M/L)$	1.5	18
Extragalactic H II regions	$10^3 (M/L)$	0.2	19
From $L = A\sigma^n$ and R_e assuming $\sigma^2 = 0.472 (GM/R_e)$: $\Sigma = 10^4 \sigma_{100}^{4-n} (L/M) (10^9/A) M_{\odot} \text{ pc}^{-2} \dagger$			
E	$10^{8.7}$	4	17
Low luminosity E	10^9	3	8
SO	10^8	6	6
From $L = A(v_{\text{rot}, 100}^{\text{max}})^n$ and R_e (assuming $\sigma = 3^{-1/2} v_{\text{rot}}^{\text{max}} \dagger$)			
Sc	$10^{9.4}$	5.2	2
Sab	$10^{8.1}$	6.4	4
Sb	$10^{8.7}$	7.0	3
Scd	$10^{8.6}$	10.0	2
Sdm	$10^{9.4}$	7.5	1.5
From (Σ, L) correlations* $\dagger\dagger$ dE $\Sigma_L = BL^{0.6}$, $\Sigma = 10^{2.9} \sigma_{100}^{3/2} (M/L)^{1/4}$ $\times (B/10^{-3.23})^{0.63} M_{\odot} \text{ pc}^{-2}$			
			8
From (Σ, L) and (L, σ) correlations* $\dagger\dagger$ cD $\Sigma_L = BL^{-1}$; $L = A\sigma^n$; $\Sigma = 10^5 \sigma_{100}^{2-n} (B/10^{14.65})^{1/2} (A/10^{8.9})$			
			8
From data			
Globular clusters $\Sigma = 10^{-0.8} (M/L) (\sigma/1 \text{ km s}^{-1})^{3.6} M_{\odot} \text{ pc}^{-2}$		1.7	26
Galaxy groups and clusters $\Sigma = 10^{0.62} \sigma_{100}^{0.8} (0.5 \text{ Mpc}/R_c)^2$		§	27

* All luminosities refer to B band.

$\dagger R_e$ is half-light radius, σ is one-dimensional velocity dispersion, Σ is mean surface density, M/R_e^2 in $M_{\odot} \text{ pc}^{-2}$, $\sigma_{100} \equiv \sigma/100 \text{ km s}^{-1}$.

$\dagger\dagger \Sigma_L$ is mean surface brightness, L/R_e^2 , in $L_{\odot} \text{ pc}^{-2}$.

§ Mean mass per galaxy assumed to be $10^{11} M_{\odot}$, R_c is the cluster core radius.

infers that

$$\Sigma \propto \sigma^{-4(n+2)(1-n)^{-1}}$$

This describes the relation of surface density to velocity dispersion for non-dissipative structure evolving from an initial power-law fluctuation spectrum; alternatively, we can regard this relation as providing the initial conditions before the onset of dissipation as it describes the evolution of both the baryonic and non-baryonic (or non-dissipative) components. Normalization is performed by comparison with the observed galaxy distribution⁴, for which the correlation function is unity on a scale of $5 h^{-1} \text{ Mpc}$ (where $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$). The remaining uncertainty is to decide what fraction the observed surface density at unit correlation length for the galaxy distribution is of the total (if indeed the luminous matter is a satisfactory tracer); this uncertainty similarly affects σ . In Fig. 1, therefore all these unknown factors are combined into the parameters Ω_b , Ω , and n ; for appropriate choices, the hierarchical clustering theory describes a unique line. Note that its location depends on both the total mass density parameter Ω (defined to be fraction of critical cosmological mass density for closure) and baryonic density parameter Ω_b , as σ depends on Ω while Σ depends on Ω_b . Hierarchical clustering evolution proceeds with increasing σ , provided $n < 1$, and towards increasing Σ , provided $n > -2$.

Next, we consider the one-dimensional collapse of an adiabatic pancake, for which

$$\Sigma = 10 (r_D/40 \text{ Mpc}) h^2 \Omega_b (1+z_p)^2 M_{\odot} \text{ pc}^{-2}.$$

symmetric collapse. Once stars begin to form, the global role of gas pressure drastically changes, and the Jeans criterion inferred from global parameters is irrelevant to continuing star formation; rather, it is dynamical pressure that is significant. Moreover, energy input from newly formed stars also considerably enhances the thermal energy of the cloud, and thereby may help regulate star formation. It seems that the global approach is too simplistic, especially if we wish to make any detailed comparison with observed systems.

A more general approach is possible, however, if we can relate the cloud surface density to that for the final system. The stellar velocity dispersion σ provides a fossilized measure of the cloud velocity dispersion at the epoch of star formation and strong dissipation. To relate Σ_{crit} to the observed Σ , we argue as follows⁵. For any system of identical clouds of surface density Σ_{crit} , the global surface density is $\Sigma = f \Sigma_{\text{crit}}$, where the surface covering or cloud overlap factor $f = t_{\text{dyn}}/t_{\text{coll}}$. Here t_{dyn} is the dynamical or collapse time for the system and t_{coll} is the cloud-cloud collision time scale. Now the luminous bulge of a galaxy will tend to be defined by cloud collisions and ensuing star formation with $t_{\text{coll}} < t_{\text{dyn}}$, simply because the region where $f < 1$ has low surface density $\Sigma < \Sigma_{\text{crit}}$. Moreover, after t_{dyn} , cloud orbits will become randomized. The resulting oblique collisions will be much less effective in compressing clouds and triggering star formation. Hence we identify the Σ for the bulge component with Σ_{crit} . The relation between Σ and σ can therefore provide us with a general framework for studying the role of dissipation in collapsing inhomogeneous star-forming clouds. Provided we consider systems of specific mass, the dynamical evolution of any substructures, which could be erased by stellar dynamical processes once dissipation stops, do not affect the (Σ, σ) correlation, because mean luminosities and half-light radii are used to derive Σ . It is also necessary that σ be interpreted as the velocity dispersion characteristic of the entire system.

The $\Sigma_{\text{crit}}(\sigma)$ relation is shown in Fig. 1 for two alternative assumptions about cloud composition, corresponding to either primordial (including hydrogen and 25% helium by mass) or solar abundances of heavy elements. The primordial composition curve illustrates that H cooling by Ly α excitation becomes ineffective below $\sigma \sim 10 \text{ km s}^{-1}$, whereas heavy elements provide fine-structure excitation degrees of freedom that allow a finite value of Σ_{crit} even at low shock velocities. The column density evaluated for solar abundances scales approximately inversely proportionally to the heavy element abundance, until the primordial values of Σ_{crit} are attained. At shock velocities above $\sim 10^3 \text{ km s}^{-1}$, free-free cooling predominates.

If we now compare the observational data with the theoretical curves, some intriguing conclusions can be drawn from inspection of Fig. 1. Galaxies and globular clusters lie in the dissipative regime, whereas galaxy clusters and groups for the most part do not. All galaxies lie in the dissipative regime, including ellipticals. The Hubble sequence reflects a sequence of end-states that have undergone varying degrees of dissipation. Giant ellipticals have not necessarily undergone extreme amounts of dissipation: for example, they could have evolved on vertical tracks of constant specific binding energy, and have undergone less dissipation than spirals. This presupposes different initial conditions for ellipticals and spirals, an option which is certainly consistent with the pancake theory of galaxy formation. According to this theory, ellipticals form in the densest pancake segments where the velocity dispersion and gas temperature may already resemble that in great galaxy clusters, whereas spirals form in cooler, more isolated regions of lower density.

Schematic evolution tracks for galaxy formation may be visualized as follows. Suppose a group of fragments or clouds destined to be a galaxy evolves as a closed self-gravitating system. The locus of a constant mass track is

$$\Sigma = 10^{2.4} (M/10^{11} M_{\odot}) \sigma_{100}^4 M_{\odot} \text{ pc}^{-2}.$$

Alternatively, suppose that an aggregate of gas clouds is gravitationally bound by non-dissipative dark matter. Then as the gas dissipates, it will maintain constant specific binding energy or

$\sigma = \text{constant}$. This should only be a transient stage, however. Once the local density is sufficiently high, self-gravity will dominate.

One might imagine that if dark potential wells of galactic mass form first, as expected in a massive gravitino-dominated universe¹⁰, the dissipation track of a characteristic baryonic element of a protogalaxy would be to evolve along the hierarchical clustering track until a mass-scale of, say, $10^{11} M_{\odot}$ was non-linear. The baryons would then evolve at constant σ until their density had increased sufficiently for baryonic self-gravity to dominate over gravitinos, and subsequent evolution would be along a track of constant mass with $\Sigma \propto \sigma_{100}^4$. The pancake fragmentation probably proceeds directly along constant mass loci, if we follow an aggregate of clouds destined to form a galaxy, as the massive neutrinos are likely to form galaxy haloes by slow infall into preexisting baryonic cores. Indeed, cosmological infall provides the most satisfactory means to date of understanding the r^{-2} density profile of isothermal haloes.

Inspection of Fig. 1 further suggests that the dwarf galaxies are located in the (Σ, σ) plane at an intermediate location between the cosmological evolutionary tracks that specify the first non-linear lumps to form and the luminous galaxies. It is tempting to infer that dwarf galaxies are the crucial link between the initial fluctuations and the luminous systems. Initially, the gas-rich dwarfs aggregated into systems of increasing surface brightness. It takes very little intergalactic gas to strip a dwarf by ram pressure¹⁴. Once there is sufficient gas, as in our halo or in groups and clusters, the gas-rich dwarfs are stripped and form dwarf spheroidals. It is these that are the fossilized remnants of galactic evolution. Dwarf spheroidals may continue to interact via dynamical friction or mergers, but the lack of dissipation greatly reduces the efficiency of any subsequent aggregation. As systems of increasing surface density and velocity dispersion are hierarchically built up by mergers of gas-rich sub-systems, stripped dwarfs remain behind as dynamical fossils. Presumably, this process accounts for the remarkable track occupied by the dEs in the (Σ, σ) plot. Ellipticals overlap disk galaxies, but extend from far lower Σ and σ to the luminous elliptical region of high Σ and σ . Globular clusters are of high central surface brightness, and may have formed in localized collapses at the same time as the galaxies. However, their location in the (Σ, σ) plane cannot be interpreted in the same way as for the other systems plotted in Fig. 1. Globular clusters have undergone considerable dynamical evolution, as indicated by their short core relaxation time scales. Much evaporative mass loss has occurred, and globular clusters cannot therefore be used to reflect any of the initial conditions relevant to formation of the galaxies.

Angular momentum has long been suspected to be the crucial parameter that determines morphological type¹¹. It is seen here that the location of a galaxy in the (Σ, σ) plane is most directly related to Hubble type and is monitoring its past evolution and degree of dissipation. Angular momentum is likely to limit dissipation for rotationally supported systems, where tidal torques of dissipating gas clouds in preexisting dark haloes can quantitatively account for the observed specific angular momenta of disks and spheroidal bulges⁹. While spirals form by slow infall of isolated aggregates of clouds into a disk, ellipticals are likely to form by the violent collapse of an inhomogeneous distribution of discrete clouds, perhaps triggered by mergers of gas-rich protospirals⁵. Ellipticals consequently can undergo considerable dissipation through cloud-cloud interactions, yet may still retain a large amount of velocity anisotropy according to dynamical collapse simulations¹² of systems containing discrete numbers of point masses (comparable with the number of clouds envisaged here). However, the actual amount of dissipation undergone by ellipticals is uncertain, especially if collapse is not limited by rotational forces. The lack of rotation of luminous ellipticals can be interpreted in two ways. Either the dissipation occurred in the absence of the dark haloes that appear to have torqued up the spirals¹³, or, as was suggested above, the degree of dissipation suffered

by many ellipticals was less than that for spirals on account of the very different initial conditions in the protogalactic cloud.

Galaxy formation may be an exceedingly slow process. The spiral galaxies gradually accumulate over the dissipation time associated with a large collection of clouds, some 10^9 years or more. Elliptical galaxies form by a more violent collapse during the initial relaxation of a galaxy cluster, but star formation in the outer regions of ellipticals may well be triggered by galaxy interactions and also extend over a characteristic time of $\sim 10^9$ yr. The correlation of morphological type with location in the (Σ, σ) plane strongly suggests that protogalactic dissipation involving gross inhomogeneities as compared with the mean observed densities accompanied by star formation had crucial roles in galactic evolution. One is tempted to envision the process of galaxy formation as a series of miniature bursts of star formation induced by cloud collisions and lasting for a billion years. Surviving clouds form dwarf galaxies, which have generally exhausted their gas supply by now, although occasional gas-rich dwarf galaxies could still be undergoing such bursts at the present time. Dwarfs are readily stripped of gas by the ram pressure of low density gaseous haloes, and dwarf spheroidals are likely to represent the resultant relics¹⁴.

Several dwarf spheroidals show indications of more recent star formation than would be inferred if all their stars formed a Hubble time ago. Gas retention and enrichment over several billions of years appears to have played a role in their evolution¹⁵. Isolated systems such as the extragalactic H II regions and the blue globular clusters around the Large Magellanic Cloud appear to have formed within the past 10^8 yr. A compact gas-rich dwarf such as I Zwicky 18 with a metallicity only 3% that of the Sun cannot have sustained its current vigorous star formation activity for more than 10^8 yr. Moreover, its current burst is almost certainly its first burst of massive star formation, otherwise excessive amounts of heavy elements would have been produced¹⁶. Objects such as this provide us with a glimpse of the faded glory that once marked the era of galaxy formation. These are the last survivors of a species that is almost extinct, the gas-rich dwarfs. The fossils remain—these are the dwarf spheroidals. Littered throughout the Universe, they are the fragments that failed to form luminous galaxies, the left-over debris that are a reminder and a signpost of galaxy formation.

I thank S.M. Fall, J. Gallagher, and P.J.E. Peebles for helpful comments, and Jean Audouze for his hospitality. This research has been partially supported by NASA and by the CNRS.

Received 15 May; accepted 21 December 1982.

1. Faber, S. & Gallagher, J. *A Rev. Astr. Astrophys.* **17**, 135–187 (1979).
2. Zel'dovich, Ya. B. *Astr. Astrophys.* **5**, 86–89 (1970).
3. Peebles, P. J. E. & Dicke, R. *Astrophys. J.* **154**, 891–908 (1968).
4. Davis, M. & Peebles, P. J. E. *Astrophys. J.* (in the press).
5. Silk, J. & Norman, C. *Astrophys. J.* **247**, 59–76 (1981).
6. Elmegreen, B. & Elmegreen, D. *Astrophys. J.* **220**, 1051–1062 (1978).
7. Silk, J. *Astrophys. J.* **211**, 638–648 (1977).
8. Rees, M. J. & Ostriker, J. P. *Mon. Not. R. astr. Soc.* **179**, 541–559 (1977).
9. Faber, S. *Astrophysical Cosmology*, (eds Bruck, H., Coyne, G. & Longair, M.), 191–234 (Specola Vaticana, Rome, 1982).
10. Blumenthal, G., Pagels, H. & Primack, J. *Nature* **299**, 37–38 (1982).
11. Sandage, A., Freeman, K. & Stokes, N. *Astrophys. J.* **160**, 831–844 (1970).
12. van Albada, T. *Mon. Not. R. astr. Soc.* **201**, 939–956 (1982).
13. Fall, S. M. *Nature* **281**, 200–202 (1979).

14. Silk, J. & Norman, C. *Astrophys. J.* **234**, 86–99 (1979).
15. Zinn, R. *Astrophys. J.* **251**, 52–60 (1981).
16. Lequeux, J., Maucherat-Joubert, M., Deharveng, J. M. & Kunth, D. *Astr. Astrophys.* **103**, 305–318 (1981).
17. Tonry, J. & Davis, M. *Astrophys. J.* **246**, 680–695 (1981).
18. Thuan, T. & Seitzer, P. *Astrophys. J.* **231**, 680–687 (1979).
19. Terlevich, R. & Melnick, J. *Mon. Not. R. astr. Soc.* **195**, 839–851 (1981).
20. Tonry, J. *Astrophys. J. Lett.* **251**, L1–L5 (1981).
21. Illingworth, G. & Kormendy, K. *Astrophys. J.* (in the press).
22. Rubin, V., Burstein, D. & Thonnard, N. *Astrophys. J. Lett.* **242**, L149–L152 (1980).
23. Roberts, M. *Astr. J.* **83**, 1026–1035 (1978).
24. Oemler, A. *Astrophys. J.* **209**, 693–709 (1976).
25. Malamuth, E. & Kirshner, R. *Astrophys. J.* **251**, 508–517 (1981).
26. Peterson, C. *Astr. J.* **81**, 617–627 (1976).
27. Bahcall, N. *Astrophys. J.* **247**, 787–791 (1981).

Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage–ligation reaction

Arthur J. Zaug, Paula J. Grabowski & Thomas R. Cech

Department of Chemistry, University of Colorado, Boulder, Colorado 80309, USA

The intervening sequence (IVS) of the Tetrahymena ribosomal RNA precursor is excised as a linear RNA molecule which subsequently cyclizes itself in a protein-independent reaction. Cyclization involves cleavage of the linear IVS RNA 15 nucleotides from its 5' end and formation of a phosphodiester bond between the new 5' phosphate and the original 3'-hydroxyl terminus of the IVS. This recombination mechanism is analogous to that by which splicing of the precursor RNA is achieved. The circular molecules appear to have no direct function in RNA splicing, and we propose the cyclization serves to prevent unwanted RNA from driving the splicing reactions backwards.

In eukaryotes, intervening sequences (IVSs) interrupt the coding regions of many genes specifying mRNA, tRNA and rRNA^{1–3}. In all cases examined, the intervening sequences are transcribed as part of a precursor RNA and are subsequently removed by a process termed RNA splicing. The splicing of one intervening sequence requires a minimum of two phosphodiester bond cleavages and one ligation. In two cases mechanisms for these cleavage–ligation events have been proposed. The splicing of yeast pre-tRNA involves independent cleavage and ligation reactions, the latter being ATP-dependent⁴. The splicing of *Tetrahymena* pre-rRNA, on the other hand, can be explained as a set of linked cleavage–ligation reactions, each of which conserves the number of phosphodiester bonds⁵. These reactions are novel in that they are mediated by the IVS RNA itself and therefore require no enzyme or other protein⁶.

By studying rRNA processing in isolated *Tetrahymena* nuclei and nucleoli, it has been shown that the IVS is excised as a unique linear molecule^{7–9}. In isolated nuclei the excised IVS is subsequently converted into a circular molecule in a reaction that occurs much more rapidly at 39 °C than at 30 °C (ref. 10). Both linear and circular forms of the IVS are produced *in vivo* in cells grown at either 30 or 39 °C (ref. 11). More recently, we have found that isolated, deproteinized linear IVS RNA undergoes autocyclization when incubated at 37–50 °C in a MgCl₂ solution^{3,6,9}.

The autocyclization reaction seems to be a useful model system for studying how RNA ligation can occur without either an enzyme or an external energy source such as ATP or GTP. We show here that the circular IVS is covalently closed by a normal 3'–5' phosphodiester linkage. The cyclization reaction involves coincident cleavage and ligation of the

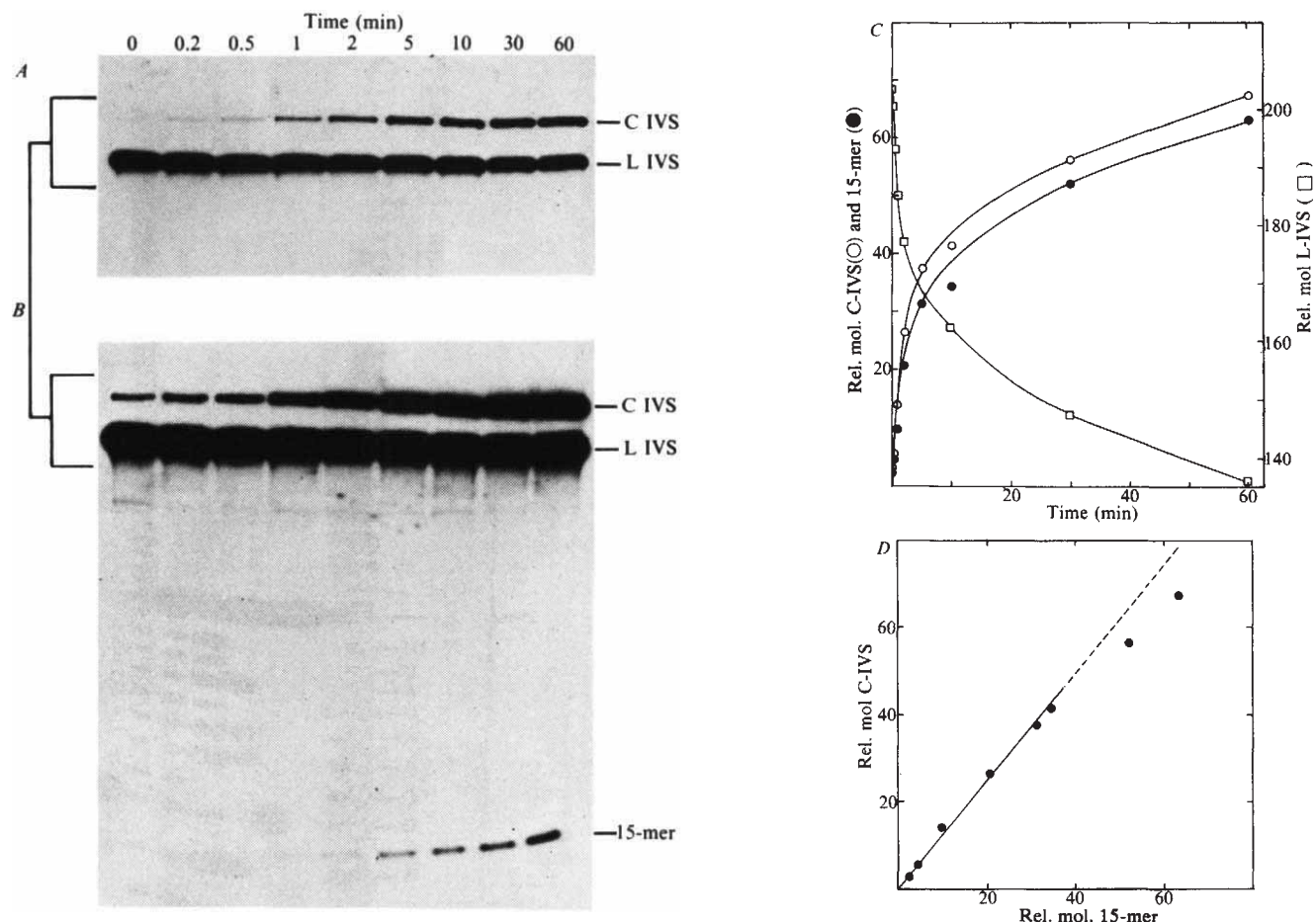


Fig. 1 Release of an oligonucleotide (15-mer) accompanies IVS RNA autocyclization. **A**, Autoradiograph of the upper portion of the 20% polyacrylamide, 8 M urea sequencing gel (4 h exposure). **B**, The gel shown in **A** exposed for 18 h to show the oligonucleotide. **C**, The radioactivity in the linear (□), circular (○) and 15-mer (●) bands of the gel was quantitated by liquid scintillation counting of the sliced gel. Relative moles of each species were calculated by dividing the number of c.p.m. by the number of A residues in the RNA (123 for the L IVS, 116 for the C IVS and 7 for the 15-mer). All values of C IVS were corrected for the small amount of that species pre-existing in the L IVS starting material. **D**, The same data are plotted to show the coincident production of the C IVS and the 15-mer.

Methods: IVS RNA was synthesized by *in vitro* transcription-splicing in isolated *Tetrahymena* nuclei⁷ at 30 °C, where little cyclization occurs¹⁰. The RNA was labelled by including [α -³²P]ATP in the transcription mixture and purified by SDS-phenol extraction. Linear IVS RNA was then isolated by sucrose gradient sedimentation and electrophoresis on a 4% polyacrylamide, 8 M urea gel¹⁰. The RNA was recovered from the gel and ethanol-precipitated. Portions of the RNA (40,000 c.p.m. each) were incubated for various lengths of time at 42 °C in 25 mM MgCl₂, 50 mM (NH₄)₂SO₄, 10 mM Tris-HCl, pH 7.5. Reactions were stopped by adding EDTA (125 mM final concentration), loaded directly onto a sequencing gel and subjected to electrophoresis for 3 h at 800 V. In **D**, at 30 and 60 min of reaction, an RNA species slightly smaller than the L IVS appeared. It was resolved from the L IVS by electrophoresis on a 4% polyacrylamide, 8 M urea gel (not shown), but not on the 20% gel shown here. Its properties were consistent with either nicked C IVS RNA or an abortive cyclization product in which the 15-mer was released without cyclization occurring. Either product would account for the decreased C IVS/15-mer ratio observed at long reaction times.

linear RNA with no net change in the number of phosphodiester bonds.

Release of an oligonucleotide accompanies IVS autocyclization

Two previous results suggested that the circular IVS did not contain all the nucleotides of the linear IVS RNA. First, when 5'-end-labelled IVS RNA was cyclized, no label was found in the circular IVS⁵. Second, at least two RNase T₁ products present in the fingerprint of the linear IVS were absent from that of the circular form⁹. Based on these observations, we analysed the products of the autocyclization reaction in such a way that any small products would be detected. Uniformly labelled linear IVS RNA was purified and incubated with MgCl₂ at 42 °C to allow cyclization, and the products were subjected to gel electrophoresis in conditions where even mononucleotides were retained on the gel. An oligonucleotide, later identified as a 15-mer (see below), was found to accumulate as IVS cyclization proceeded (Fig. 1A, B). The oligonucleotide and the circular IVS were produced in equimolar amounts from the earliest reaction times tested (Fig. 1C, D). We have also

looked at autocyclization in the presence of low concentrations of formamide, with Mn²⁺ substituted for Mg²⁺, and following treatment of the linear IVS RNA with bacterial alkaline phosphatase to remove its 5'-terminal phosphate; none of these perturbations uncoupled IVS cyclization from production of the oligonucleotide (A.J.Z. and T.R.C., unpublished results). We conclude that IVS cyclization is concomitant with and probably dependent on the release of this oligonucleotide.

Structure of the oligonucleotide

The products of autocyclization of uniformly labelled IVS RNA were examined by electrophoresis on a 20% polyacrylamide sequencing gel (Fig. 2A). The oligonucleotide co-migrated with the 16th band on a partial alkaline hydrolysis ladder, consistent with it being a linear RNA with the structure pN₁₆p or pN₁₅OH. The oligonucleotide was recovered from such a gel and subjected to RNase T₁ fingerprinting (data not shown). It was found to contain three T₁ fragments: pGp, which must arise from the 5' end of the IVS RNA⁹; the T₁ hexanucleotide that follows pGp in the sequence (spot 2 of ref. 9); and a new product of medium size. The identification of the RNase T₁ product pGp

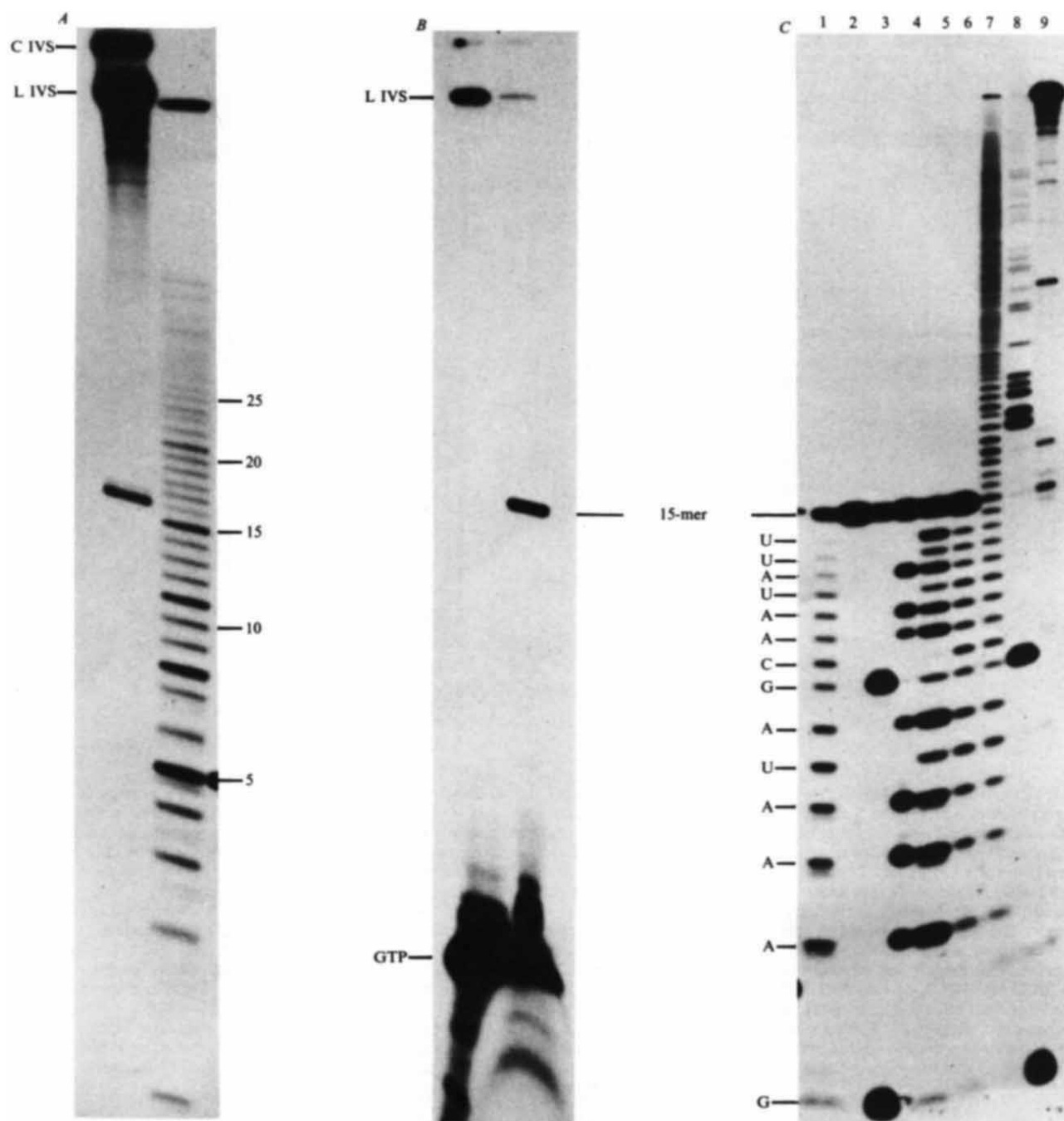


Fig. 2 The oligonucleotide contains the first 15 bases of the linear IVS RNA. **A**, Autoradiogram of uniformly-labelled IVS RNA after autocyclization (left lane) analysed by gel electrophoresis with oligonucleotide markers (right lane). **B**, Autoradiogram of 5' end-labelled IVS RNA incubated in control (left lane) and autocyclization (right lane) conditions. **C**, Lanes 1-6, the 15-mer purified from the gel shown in **B**; lanes 7-9, the linear IVS RNA, were analysed by the enzymatic sequencing gel method^{12,13}. Lanes 1, 6, 7, partial alkaline hydrolysis; lanes 2, 9, RNA incubated in the absence of enzymes. RNA was also subjected to limited digestion by RNase T₁ (lanes 3, 8), RNase U₂ (lane 4) and RNase Phy M (lane 5). Note that the 15-mer (pN₁₅OH) co-migrates with the 16th partial digestion product (pN₁₆p) of the linear IVS RNA.

Methods: **A** left lane: linear IVS RNA was produced by *in vitro* transcription-splicing in isolated *Tetrahymena* nuclei. The RNA was uniformly labelled with both [α -³²P]ATP and [α -³²P]GTP during transcription. The linear IVS RNA was purified and incubated for 20 min in autocyclization conditions as described in Fig. 1 legend. After electrophoresis on a 20% polyacrylamide, 8 M urea sequencing gel, the autoradiograph shown here was obtained. Right lane: unlabelled linear IVS RNA was purified, treated with bacterial alkaline phosphatase and end-labelled with polynucleotide kinase as described by A.J.Z. and T.R.C.⁹. It was then subjected to limited alkaline hydrolysis to provide size markers for the oligonucleotide in the left lane. Note that the RNAs in both lanes have 5'-monophosphate ends. **B**, Linear IVS RNA was produced by *in vitro* splicing of plasmid DNA transcripts. Supercoiled pIVS11 DNA, which has the IVS-containing 1.6-kilobase (kb) *Hind*III fragment of *T. thermophila* rDNA inserted downstream from a *lac* UV5 promoter, was transcribed with purified *E. coli* RNA polymerase in non-splicing conditions as described by Kruger *et al.*⁶. No ³²P-labelled nucleoside triphosphates were present during the transcription. The RNA was purified and incubated for 30 min at 30 °C in splicing conditions⁶ using [α -³²P]GTP as the cofactor. This procedure results in labelling of the excised IVS RNA at its 5' end and retention of all three phosphates from the GTP^{5,6}. Left lane: a portion of the RNA was incubated for 15 min at 40 °C in the absence of MgCl₂ as a control for the autocyclization reaction. Sequencing gel electrophoresis and autoradiography revealed labelled linear IVS RNA and unincorporated GTP. Right lane: a portion of the RNA was incubated for 15 min at 40 °C in the autocyclization conditions described in Fig. 1 legend. Most of the radioactivity originally present in the linear IVS RNA appeared in the 15-mer. The unincorporated GTP was still present.

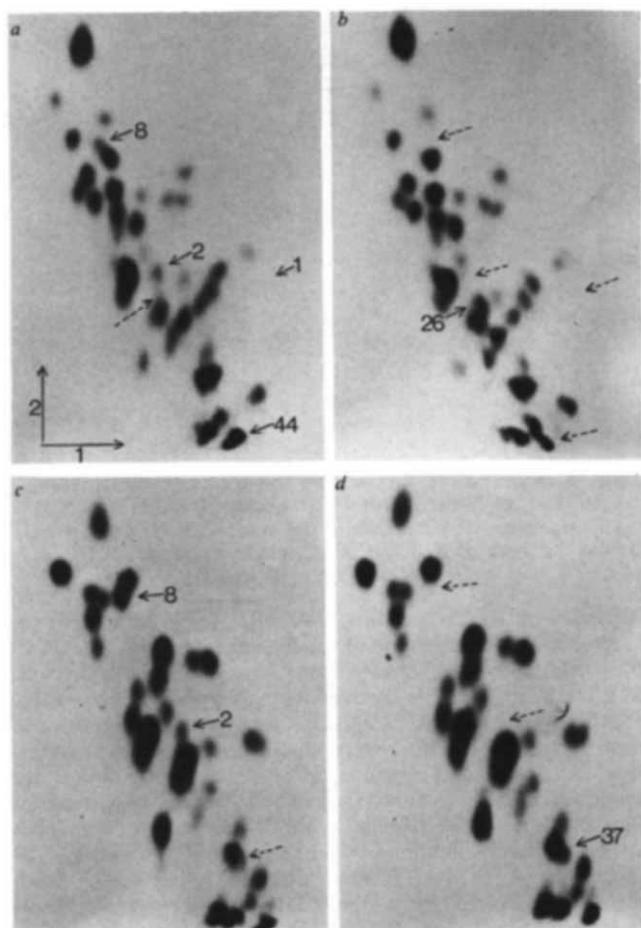


Fig. 4 RNase T₁ fingerprint analysis of linear and circular IVS RNAs (a and c). Linear IVS RNA produced by transcription-splicing in isolated *Tetrahymena* nuclei. One transcription reaction (a) contained [α -³²P]ATP as the only labelled nucleotide, and another (c) contained only [α -³²P]CTP. The linear IVS RNA from each reaction was purified as described in Fig. 1 legend. A portion of each sample was then incubated in autocyclization conditions and subjected to electrophoresis on a 4% polyacrylamide, 8 M urea gel. Using an autoradiograph as a guide, the gel was sliced and the circular IVS RNA recovered by crushing and soaking. The linear (a, c) and circular (b, d) IVS RNAs were digested to completion with RNase T₁ and subjected to fingerprinting as described previously⁹. First dimension was electrophoresis on cellulose acetate at pH 3.5, second dimension was PEI chromatography at pH 3.5. —→, Oligonucleotides unique to either the linear or the circular IVS RNA; - - -→, spots missing or decreased in intensity in one of the two forms. Oligonucleotides were identified as described by A.J.Z. and T.R.C.⁹; a few tentative identifications have been resolved by this work (Table 1). (1) pGp, the 5' terminal oligonucleotide, visible on longer exposure of fingerprint a. (2) AAAUAGp, which follows pGp in the sequence. (8) UACUCG-OH, the 3'-terminal oligonucleotide. Other numbered spots are identified in Table 1.

cyclized IVS. These are the sequences that make up the 15-mer: spot 1 (pGp, present only in the A-labelled RNA fingerprint but difficult to see in the photograph in Fig. 4a), spot 2 (AAAUAG, easily seen in Fig. 4a and absent from Fig. 4b), and a portion of spot 44. The RNase T₂ data in Table 1 show clearly that one of the two oligonucleotides in spot 44, the one that follows AAAUAG in the linear IVS (Fig. 3), is lost on cyclization. The two new oligonucleotides in the circular IVS that were predicted to border the cyclization junction were both identified. The UACUCGp oligonucleotide co-migrates with two others in spot 26C, but most of this background is eliminated in the ATP-labelled RNA (compare Fig. 4a, b). The occurrence of a molar amount of labelled Gp in the RNase T₂ digest of spot 26C RNA (Table 1) shows that cyclization invol-

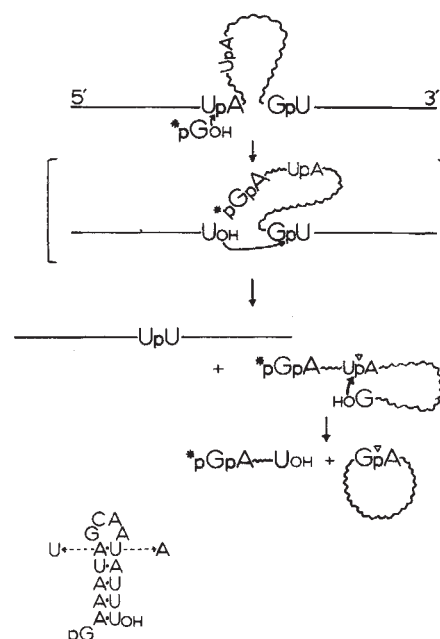


Fig. 5 Phosphoester transfer model for pre-rRNA splicing and IVS cyclization. —, Exons; ~, IVS; *, labelled guanosine cofactor, in this example 5'-GMP; ▽, phosphate that closes the circle. Square brackets denote a postulated intermediate that has not been isolated. Bottom: possible secondary structure of the 15-mer released during cyclization; - - -→, sequence differences in the IVS of *T. pigmentosa* 6 UM²⁶.

ves formation of a phosphodiester bond between this G and an A, presumably A₁₆. The identification of ACCUUGp in spot 37C (CTP label, Table 1) confirms that cyclization involves interjection of a G to the left of A₁₆.

We conclude that IVS RNA cyclization takes place by the cleavage-ligation reaction shown schematically in Fig. 3. The linkage between G₄₁₄ and A₁₆ that forms the circle is cleaved by RNase T₁ to produce the oligonucleotides in spots 26C and 37C. The linkage must therefore be a normal 3'-5' phosphodiester bond. We also note that the bond in the linear IVS that is interrupted by the cyclization reaction, the bond between U₁₅ and A₁₆, is a normal phosphodiester bond. It is susceptible to cleavage by RNase T₂ (Table 1, ATP-labelled RNA in spot 44L was completely digested) and by alkali (Fig. 2C).

Phosphoester transfer, a novel mechanism for RNA ligation

We have shown that cyclization of the *Tetrahymena* ribosomal IVS is a cleavage-ligation reaction. The linear RNA is cleaved 15 nucleotides from its 5' end, and the A residue at the cleavage site is joined to the G at the 3' terminus by a normal phosphodiester bond. The phosphate that forms the linkage comes from the internal position between residues 15 and 16 in the linear molecule. Cyclization requires no external energy source such as ATP or GTP. We therefore suggest that cleavage of the phosphodiester bond is providing the energy for the ligation event and that the two reactions are concerted. Consistent with such a phosphoester transfer mechanism, we have found that linear IVS RNA molecules lacking the 15-mer (either broken circular IVS molecules or abortive cyclization products) are unable to cyclize (P.J.G., unpublished results).

The steps involved in cyclization of the IVS are detailed in Fig. 5, together with the mechanism proposed for pre-rRNA splicing^{8,9}. The splicing and cyclization reactions seem to occur by the same phosphoester transfer mechanism. All cleavage reactions leave 5'-phosphate and 3'-hydroxyl termini. All ligation steps involve joining of the newly created 5' phosphate to a different 3'-hydroxyl terminus, resulting in strand transfer. IVS cyclization also resembles the first step in pre-rRNA

splicing with respect to the nucleotides involved. Both reactions involve cleavage of the sequence UpA and attack by a G-OH, forming a GpA linkage. The two reaction sites are not equivalent, however, because the initial splicing reaction does not occur to any detectable extent at the cyclization site⁵. In any case, the UA 'consensus sequence' is so short that it must be only one of several RNA structural features that determine the site of a phosphoester transfer reaction.

The mechanism of cyclization of the *Tetrahymena* IVS is quite different from that involved in cyclization of RNA by T4 RNA ligase. Ligase catalyses end-to-end cyclization of RNA chains^{15,16}. The 5'-phosphoryl end of the RNA is activated by adenylation, and attack by the 3'-hydroxyl end results in formation of a normal phosphodiester bond^{17,18}. Recently, an RNA ligase from wheat germ has also been found to cyclize a variety of RNAs that have a 2', 3'-cyclic phosphate terminus (refs 19, 20 and P. Gegenheimer, C. Peebles, C. Greer and J. Abelson, manuscript in preparation). Although the requirement for the 3'-terminal cyclic phosphate defines a new class of RNA ligase, the wheat germ enzyme resembles T4 RNA ligase in that it requires ATP, adenylates 5' termini and catalyses end-to-end cyclization.

Although the mechanism of cyclization of the *Tetrahymena* IVS RNA does not resemble these RNA ligase mechanisms, it is similar to the ATP-independent breakage-reunion process catalysed by DNA topoisomerases²¹. In the replication of Φ X174 DNA, single-stranded viral DNA is cleaved and cyclized by the A or A* protein²²⁻²⁴ in a phosphoester transfer reaction that is very similar to that described for IVS cyclization. DNA recombination also requires a formally similar process of cleavage, strand switching and rejoining.

Unlike these DNA breakage-reunion processes, pre-rRNA splicing and IVS cyclization require no enzyme. The cleavage-ligation activity is intrinsic to the structure of the RNA⁶. IVS cyclization provides a well defined model system for future studies aimed at understanding how an RNA molecule can break and rejoin specific covalent bonds.

Although we do not know what sequences or structural features of the RNA are required for reactivity, the first 15 nucleotides of the IVS comprise one region of interest. This sequence can be folded into a hairpin structure with a 5-base pair stem and a 4-base loop (Fig. 5). Cleavage of the pre-rRNA and G addition take place on one side of the base of the stem, while IVS cyclization occurs on the other side. This is predicted to be a rather unstable structure; $\Delta G^\circ = -1.5 \text{ kcal mol}^{-1}$, calculated as described in ref. 25 using a revised value of $+4.5 \text{ kcal mol}^{-1}$ for formation of the loop (I. Tinoco Jr, personal communication). Although the sequences of the IVSs of *T. thermophila* and *Tetrahymena pigmentosa* 6UM differ in two positions in this region^{14,26}, the changes are such that the potential for base-pairing is preserved (Fig. 5). Such covariant changes have proven to be powerful predictors of secondary structure in the cases of tRNA, 5S rRNA and other rRNAs^{27,28}. Therefore, although more sequences are needed for comparison, we

tentatively conclude that the 15-mer has a hairpin secondary structure at some stage in the reaction. Site-specific mutagenesis should allow us to determine what role (if any) this marginally stable hairpin might have in IVS excision and/or cyclization.

Why circular IVS RNA?

The low steady-state concentration of circular IVS RNA present in *Tetrahymena* *in vivo* is consistent with a very short half life for this molecule, of the order of 6 s (ref. 11). Some strains of *Tetrahymena* have no excised IVS RNA because they have no IVS in their rDNA^{29,30} or encoded elsewhere in their genome³¹. These facts discourage us from postulating a function for the circular IVS RNA. The complete separability of the IVS excision and cyclization reactions precludes the cyclization process from having a direct function in the splicing process *in vitro* or presumably *in vivo*. Instead, we propose that IVS cyclization serves to destroy the non-essential IVS product of rRNA splicing, thereby giving a forward direction to an otherwise reversible reaction. What would be the consequence of the reverse reaction? The reinsertion of the IVS into already spliced rRNA would slow the process of rRNA maturation. Another possible reverse reaction, insertion of the excised IVS into a different site in an RNA molecule (transposition), would have more detrimental consequences. Cyclization could inactivate the IVS by cleaving it into two fragments and tying up its 3' end in a phosphodiester bond, thereby preventing any reverse splicing reaction.

How general is the process of cyclization of intervening sequences or the ligation of any RNA by a breakage-reunion mechanism? Some yeast mitochondrial mRNA intervening sequences are circular^{32,33}. It is unknown, however, whether the circles are the direct or indirect products of splicing, whether they are closed by a phosphodiester bond, or whether any nucleotides are released during cyclization. Intervening sequences excised from nuclear pre-tRNA in yeast are linear molecules³⁴. In very few other cases has any form of an excised IVS, linear or circular, been detected (see ref. 11); any conclusion about the generality of IVS cyclization would therefore be premature.

Viroids are single-stranded, covalently closed circles of RNA³⁵ approximately the same size as the *Tetrahymena* IVS. Branch *et al.*³⁶ have described unit-length and greater-than-unit-length linear (+)-strand RNAs that seem to be intermediates in viroid replication. It will be interesting to determine whether formation of the mature, circular viroid involves end-to-end joining of a linear intermediate or, like *Tetrahymena* IVS RNA cyclization, occurs via cleavage-ligation at an internal point in a linear molecule.

We thank Olke Uhlenbeck, Nick Cozzarelli, Hugh Robertson and Jim Dahlberg for helpful discussions and Kelly Kruger for providing pIVS11 DNA. This research was supported by American Cancer Society grant NP-374 and NIH grant GM28039 to T.R.C. T.R.C. was supported by a Research Career Development Award from the NCI, DHHS.

Received 2 November; accepted 29 November 1982.

- Abelson, J. A. *Rev. Biochem.* **48**, 1035-1069 (1979).
- Breathnach, R. & Chambon, P. A. *Rev. Biochem.* **50**, 349-383 (1981).
- Cech, T. R., Zaug, A. J., Grabowski, P. J. & Brehm, S. L. *The Cell Nucleolus* Vol. 10 (eds Busch, H. & Rothblum, L.) 171-204 (Academic, New York, 1982).
- Peebles, C. L., Ogden, R. C., Knapp, G. & Abelson, J. *Cell* **18**, 27-35 (1979).
- Cech, T. R., Zaug, A. J. & Grabowski, P. J. *Cell* **27**, 487-496 (1981).
- Kruger, K. *et al. Cell* **31**, 147-157 (1982).
- Zaug, A. J. & Cech, T. R. *Cell* **19**, 331-338 (1980).
- Carin, M. *et al. Nucleic Acids Res.* **8**, 5551-5566 (1980).
- Zaug, A. J. & Cech, T. R. *Nucleic Acids Res.* **10**, 2823-2838 (1982).
- Grabowski, P. J., Zaug, A. J. & Cech, T. R. *Cell* **23**, 467-476 (1981).
- Brehm, S. L. & Cech, T. R. *Biochemistry* (in the press).
- Donis-Keller, H., Maxam, A. M. & Gilbert, W. *Nucleic Acids Res.* **4**, 2527-2538 (1977).
- Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. & Guillely, H. *Nature* **269**, 833-836 (1977).
- Kan, N. C. & Gall, J. G. *Nucleic Acids Res.* **10**, 2809-2822 (1982).
- Silber, R., Malathi, V. G. & Hurwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3009-3013 (1972).
- Bruce, A. G. & Uhlenbeck, O. C. *Nucleic Acids Res.* **5**, 3665-3677 (1978).
- Higgins, N. P. & Cozzarelli, N. R. *Meth. Enzym.* **68**, 50-71 (1979).
- Uhlenbeck, O. C. & Gumpert, R. I. *Enzymes* **15**, 31-58 (1982).

- Konarska, M., Filipowicz, W., Domdey, H. & Gross, H. L. *Nature* **293**, 112-116 (1981).
- Konarska, M., Filipowicz, W. & Gross, H. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1474-1478 (1982).
- Gellert, M. A. *Rev. Biochem.* **50**, 879-910 (1981).
- Eisenberg, S., Scott, J. F. & Kornberg, A. *Cold Spring Harb. Symp. quant. Biol.* **43**, 295-302 (1978).
- Sumida-Yasumoto, C. *et al. Cold Spring Harb. Symp. quant. Biol.* **43**, 311-329 (1978).
- van der Ende, A., Langeveld, S. A., Teertstra, R., van Arkel, G. A. & Weisbeek, P. J. *Nucleic Acids Res.* **9**, 2037-2053 (1981).
- Tinoco, I. Jr *et al. Nature new Biol.* **246**, 40-41 (1973).
- Wild, M. A. & Sommer, R. *Nature* **283**, 693-694 (1980).
- Fox, G. & Woese, C. R. *Nature* **256**, 505-507 (1975).
- Noller, H. F. & Woese, C. R. *Science* **212**, 403-410 (1981).
- Wild, M. A. & Gall, J. G. *Cell* **16**, 565-573 (1979).
- Din, N. & Engberg, J. *J. molec. Biol.* **134**, 555-574 (1979).
- Wild, M. A. thesis, Yale Univ. (1980).
- Halbreich, A., Pajot, P., Foucher, M., Grandchamp, C. & Slonimski, P. *Cell* **19**, 321-329 (1980).
- Armstrong, A. C., VanOmmen, G.-J. B., Grivell, L. A., VanBruggen, E. F. J. & Borst, P. *Cell* **19**, 313-319 (1980).
- Knapp, G., Ogden, R. C., Peebles, C. L. & Abelson, J. *Cell* **18**, 37-45 (1979).
- Gross, H. J. *et al. Nature* **273**, 203-208 (1978).
- Branch, A. D. *et al. Science* **217**, 1147-1149 (1982).

How galaxies acquire their neutrino haloes

J. Richard Bond

Physics Department, Stanford University, Stanford, California 94305, USA

Alex S. Szalay & Simon D. M. White

Astronomy Department, University of California, Berkeley, California 94720, USA

One-dimensional simulations of the nonlinear growth of structure in a universe dominated by a population of nonrelativistic collisionless particles such as massive neutrinos show that a subpopulation of slowly moving particles exists within the 'pancakes' that form¹. These particles can cluster in a low velocity condensate around any seed perturbation which may be present. The schematic calculation of this aggregation presented here suggests that the properties of neutrino clusters depend only weakly on seed mass but substantially on seed separation. Their mass and velocity dispersion may be quite comparable with the values inferred for the haloes of 'dark' matter surrounding real galaxies.

Reports of a non-zero neutrino rest mass^{2,3} have encouraged speculation that the mass budget of the Universe may be dominated by such particles. Neutrinos would then be the 'dark' matter which may make up massive and as yet unseen haloes around many galaxies⁴. No small-scale fluctuations in a cosmological population of neutrinos can survive early epochs when the neutrino velocity dispersion was high^{5,6}. Detailed calculations^{7,8} show that fluctuations are suppressed on scales smaller than a co-moving wavelength with present size $\lambda_0 = 41 m_{30}^{-1}$ Mpc, where m_{30} is the neutrino mass in units of 30 eV

and one neutrino species is assumed to carry all of the mass. This is the scale of the first objects to separate from the overall expansion in a neutrino-dominated universe. The formation of such objects must have occurred quite recently ($z < 10$) to avoid overly large temperature fluctuations in the microwave background^{5,6}, and is expected to involve an essentially one-dimensional collapse to a sheet-like 'pancake' configuration⁹. Several authors^{1,11} have, therefore, idealized this collapse by one-dimensional calculations. *N*-body simulations of this type are illustrated in Fig. 1. The motion of 6,000 neutrino sheets per wavelength was followed self-consistently by evaluating the force on each by direct summation. Such calculations run rapidly on a small computer and conserve energy to better than 1 part in 10^4 (the degree to which the Layzer-Irvine equation is satisfied). The initial conditions were those of the Zeldovich approximation^{9,11} which is exact in one dimension for cold particles. Thermal velocities were randomly chosen from the background neutrino distribution⁵ projected onto one dimension and adiabatically compressed. These were then added to the Hubble flow and perturbation-induced velocities in setting up the initial velocities of the neutrino sheets. Figure 1*a, b* shows how the nonlinear growth of an initially sinusoidal perturbation leads to a distribution which is strongly concentrated to the pancake plane and contains well-defined interpenetrating particle streams. The three-dimensional 'thermal' velocity dispersion of the background neutrinos is $6 m_{30}^{-1} (1+z)$ km s⁻¹, which is small by the nonlinear regime and has little effect on such phase-space diagrams. The dispersion in the rather 'hot' example of Fig. 1 is at the limit of plausibility for the real Universe. The one-dimensional, multi-stream structure of the velocity distribution gives rise to many low velocity particles which are available for capture onto any seed condensations which may form. In Fig. 1*c, d* we show how the fraction of the total neutrino population moving more slowly than any given velocity depends on distance from the pancake plane. Such distributions depend only weakly on collapse epoch, and scarcely at all on either the residual neutrino velocity

Fig. 1 Simulations of the evolution of a one-dimensional sinusoidal perturbation in a neutrino-dominated Einstein-de Sitter universe. *a* and *b* show the distribution of particles in phase-space at expansion factors of 0.34 and 4.7 relative to the time, t_c , when neutrino streams first interpenetrate ('pancaking'). Peculiar velocity with respect to the expansion is plotted against x , the position expressed in units of half the wavelength, $\lambda(t)/2$. The velocity normalization is $\sigma_n = \lambda(t_c)/2t_c = 1.732 m_{30}^{-1/2} (1+z_c)^{1/2}$ km s⁻¹ where z_c is the redshift corresponding to t_c . (Note that for a flat universe dominated by a single neutrino species $H_0 = 56 m_{30}^{1/2}$ km s⁻¹ Mpc⁻¹.) The three-dimensional thermal dispersion is $0.035\sigma_n$ in *a* but decays rapidly as the universe expands, and is only $0.0026\sigma_n$ in *b* which accounts for the large decrease in fuzziness between *a* and *b*. *c* and *d* show the cumulative total velocity distribution within various distances of the central plane (normalized as in *a*) and are divided by the total number of sheets; they give the distributions now for collapse epochs $1+z_c = 3$ and 4.7 respectively, or equivalently for expansion factors relative to collapse, $(1+z_c)/(1+z)$, of 3 and 4.7. Note the large number of low velocity neutrinos near the pancake which cool off as the universe expands.

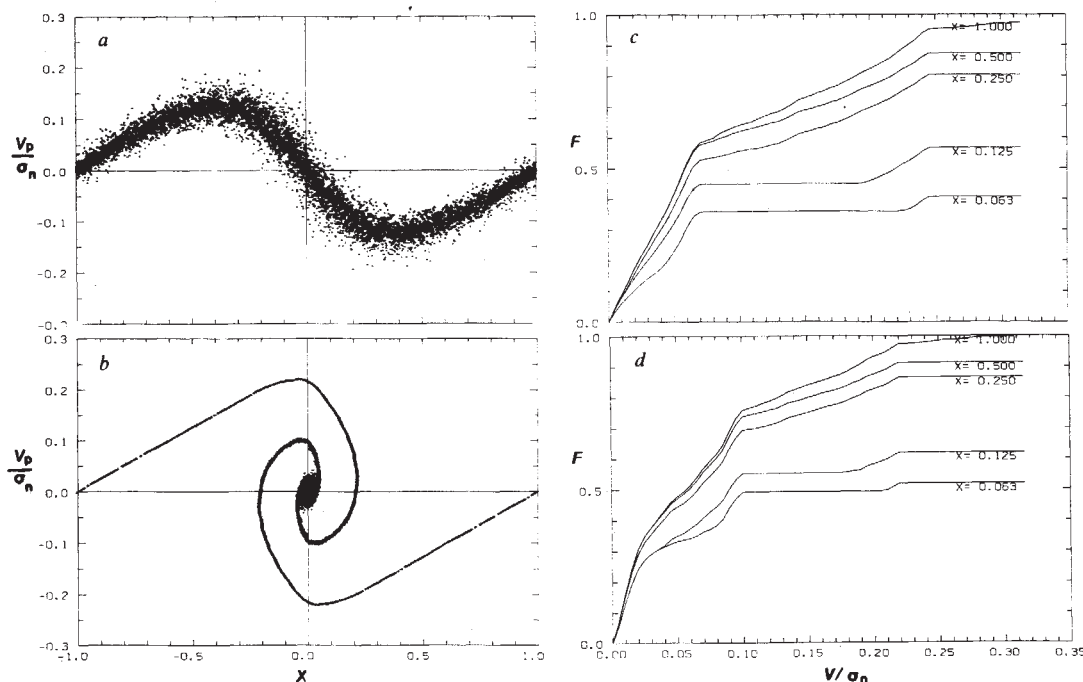
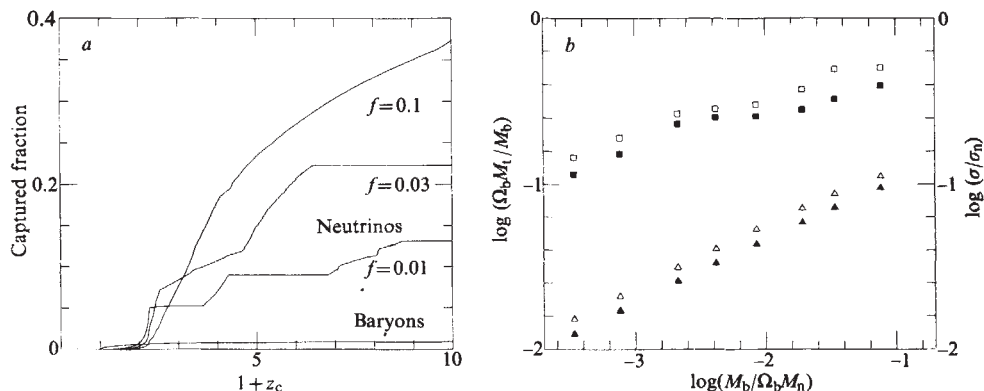


Fig. 2 Capture of neutrinos by small baryon seeds. *a*, The fraction of the total mass of the universe which has been captured as a function of expansion factor for $\Omega_b = 0.01$ and for three different values of the seed spacing parameter f . The condensations rapidly become neutrino-dominated. *b*, Results are plotted at $1+z_c=6$ for various values of f and Ω_b . The ratio of total-to-baryonic mass (squares) and the three-dimensional velocity dispersion (triangles) in individual haloes are given as a function of their baryonic mass, M_b . Filled symbols correspond to $\Omega_b = 0.01$ and open symbols to $\Omega_b = 0.03$. The normalizations are $M_n = 4.8 \times 10^{15} m_{30}^{-2} M_\odot$ and $\sigma_n = 4,235 m_{30}^{-1/2} \text{ km s}^{-1}$. Note the smooth trends of these quantities with seed spacing ($f \propto (M_b/\Omega_b)^{1/2}$) and their weak dependence on seed mass ($\propto \Omega_b$ at fixed f).



dispersion or the presence or absence of a thin layer of shocked baryonic matter in the pancake plane (except at very low velocities, $< 50 \text{ km s}^{-1}$). The phase-space distribution shown in Fig. 1 cannot apply to baryonic material which must shock before it hits the central plane.

In suitable conditions, some fraction of this material will radiate its energy and form a thin cold layer which may then fragment¹¹. Such fragmentation must occur if this picture is to produce the Universe we see, but the physics which governs it is both complex and ill understood; we assume that fragmentation does indeed take place and enquire whether low velocity neutrinos can condense onto baryon seeds. Because we wish to grow massive haloes around small seeds, the condensation process can only be handled correctly by a fully self-consistent three-dimensional calculation. Rather than attempt such a treatment we here present a schematic calculation of neutrino capture based on our one-dimensional evolution program. We designate some small fraction Ω_b of each neutrino sheet as baryonic matter. As a sheet crosses the pancake plane for the first time these baryons are assumed to shock, to separate from the neutrinos and to collapse onto condensation sites distributed in the plane with surface number density $[\pi(f\lambda)^2]^{-1}$. In addition, every time a neutrino sheet crosses the pancake plane we calculate the fraction of its area which is sufficiently close to a condensation centre for its binding energy to be negative; this fraction is assumed to be captured and the sheet mass is decreased and the mean mass and binding energy of the haloes are augmented by appropriate amounts. We model the gravitational potential of a halo by assuming its density-radius relation to be an inverse square law within a cut-off radius which can be derived using its total mass and the specific binding energy of previously captured neutrinos. We approximate the haloes by a uniform layer in the pancake plane when calculating the force they exert on other neutrino sheets.

Figure 2 gives some results of using the above procedure in one-dimensional simulations like those of Fig. 1. The growth of neutrino haloes is only weakly dependent on the baryon fraction, because captured neutrinos soon dominate the potential of the lumps and thereafter regulate further capture themselves. Capture efficiency depends strongly on the spacing of the seeds, however, as this determines how much of the captured mass is associated with each halo, and so how effectively it can catch more neutrinos.

If all bright galaxies were to form in pancakes, then to reproduce their observed density (M. Davis, personal communication) we require $(\pi f^2 \lambda^3)^{-1} = 0.01 h_{100}^3 \text{ Mpc}^{-3}$ hence $f = 0.05 m_{30}^{3/4}$. Values of the dimensionless spacing parameter of this order are sufficient to ensure the capture of a substantial fraction of the low velocity neutrinos. The specific binding energy of captured neutrinos also agrees reasonably well with the specific energy inferred for the dark matter in galaxy haloes. For example, the $\Omega_b = 0.01$ case in Fig. 2 is well fitted by the

three-dimensional velocity dispersion $\sigma \approx 10^2 m_{30}^{0.26} (M_b/10^{11} M_\odot)^{0.38} \text{ km s}^{-1}$ and only approximately fit by $M_t \sim 2 \times 10^{12} m_{30}^{0.4} (M_b/10^{11} M_\odot)^{1.2} M_\odot$ relating total mass to that in baryons. Lower velocity dispersions are found for $\Omega_b = 0.03$ and higher ones for smaller Ω_b due to the inverse dependence of seed spacing on Ω_b at fixed M_b . Here, Ω_b should be interpreted as the fraction of baryons that have cooled sufficiently to become galaxy seeds. It can be substantially less than the total baryon fraction and may lead to the $\sim 200 \text{ km s}^{-1}$ dispersions characteristic of bright spiral galaxies. The one-at-a-time addition of neutrinos from a low velocity distribution onto already existing halo seeds will give a less concentrated halo than the $\sim r^{-3}$ -density profile found when the collisionless matter relaxes violently in a single collapse event^{12,13}. Note that our models predict that the ratio of total to baryonic mass and the overall velocity dispersion of galaxy-halo systems should both increase with baryonic mass, and thus, presumably, with luminosity. Such dependences are indeed observed⁴.

Our results were obtained for $\Omega = 1$; lowering Ω will hardly change the nonlinear evolution along the collapse axis, but will result in more rapid expansion in the transverse directions which may aid in lowering the neutrino velocities to yield slightly larger capture fractions and specific binding energies.

Our schematic calculations appear to be supported by the fully self-consistent two-dimensional simulations of Melott¹⁴, although he gave little quantitative data. The pancaking and low-velocity condensation processes involve little dilution of the initial neutrino phase-space density. Nevertheless, phase-space constraints imply that the density of a neutrino halo cannot continue to rise as the inverse square of radial distance within 10 kpc of the centre of a galaxy¹⁵. In addition, in the real world pancakes are unlikely to remain one dimensional over time scales in excess of their initial collapse time. This should be borne in mind when evaluating the results of our own and similar work.

This work was partially supported by the NSF under grant AST-8114715 and by DOE contract AT03-82ER40069.

Received 27 July; accepted 3 December 1982.

- Melott, A. L. *Phys. Rev. Lett.* **48**, 894-896 (1982).
- Lyubimov, V. A., Novikov, E. G., Nozik, V. Z., Tretyakov, E. F. & Kozik, V. S. *Phys. Lett.* **B94**, 266-268 (1980).
- Reines, F., Sobel, S. W. & Pasierb, E. *Phys. Rev. Lett.* **45**, 1307-1311 (1980).
- Faber, S. M. & Gallagher, J. S. A. *Rev. Astr. Astrophys.* **17**, 135-187 (1979).
- Bond, J. R., Efstathiou, G. & Silk, J. I. *Phys. Rev. Lett.* **45**, 1980-1984 (1980).
- Doroshkevich, A. G., Khlopov, M. Yu., Sunyaev, R. A., Szalay, A. S. & Zel'dovich, Ya. B. *Ann. N.Y. Acad. Sci.* **375**, 32-42 (1980).
- Peebles, P. J. E. *Astrophys. J.* **258**, 415-424 (1982).
- Bond, J. R. & Szalay, A. S. *Astrophys. J.* (in the press).
- Zel'dovich, Ya. B. *Astr. Astrophys.* **5**, 84-89 (1970).
- Doroshkevich, A. G. et al. *Mon. Not. R. astr. Soc.* **192**, 321-327 (1980).
- Sunyaev, R. A. & Zel'dovich, Ya. B. *Astr. Astrophys.* **20**, 189-200 (1972).
- White, S. D. M. *Mon. Not. R. astr. Soc.* **189**, 831-852 (1979).
- van Albada, T. *Mon. Not. R. astr. Soc.* **201**, 939-955 (1982).
- Melott, A. L. *Nature* **296**, 721-723 (1982).
- Tremaine, S. D. & Gunn, J. E. *Phys. Rev. Lett.* **42**, 407-410 (1979).

Trail of the Crab progenitor star

Roger D. Blandford, Charles F. Kennel
& Christopher F. McKee

Institute for Theoretical Physics, University of California,
Santa Barbara, California 93106, USA

Jeremiah P. Ostriker

Princeton University Observatory, Princeton, New Jersey 08540, USA

Gull and Fesen^{1,2} have shown that the 'jet', observed by van den Bergh³ to protrude from the northern rim of the Crab Nebula, appears to be an edge-brightened cylinder with astonishingly straight sides. This faint feature, which was observed in both oxygen and hydrogen emission lines, has a radial velocity within 150 km s^{-1} of that of the nebula. It has been suggested⁴ that the jet arises from an instability in the outskirts of the visible nebula. We propose here that the jet is a tubular trail of material lost by the pre-supernova star during its red giant phase.

Recent optical observation's support Chevalier's⁵ view that the Crab Nebula was formed in a Type II supernova. The radio and optical nebula appears to be surrounded by a shell of diffuse H α emission⁶ with radial velocities relative to the nebula as large as $7,000 \text{ km s}^{-1}$. Extended soft X-ray emission may also have been detected^{8,9}. These observations are consistent with several solar masses of ejecta expanding roughly ballistically into the low density, hot phase of the interstellar medium¹⁰.

It seems unlikely that such a delicate feature as the observed jet can have survived the blast wave that precedes the supernova ejecta. From the age of the remnant, we estimate that the radius of the blast wave is roughly 7 pc, consistent with the reported dimensions of the H α shell⁶. The jet is located at a projected distance of ~ 2 pc from the centre of nebular expansion (~ 0.1 pc to the east of the pulsar), and appears to connect directly to the nebula. However, there appears to be no structure within the nebula that connects to the jet^{1,2}. If the radius of the blast wave is 7 pc, then the trail we propose must be inclined to the line of sight at an angle of $\sim 15^\circ$, in order that the portion of the trail that was destroyed by the supernova remnant be obscured by the intense emission from the Nebula. None of our estimates below depends critically on this angle.

The progenitor of the Crab pulsar presumably went through a red giant phase before its explosion. Red giants lose mass at a rate $\dot{M} = 10^{-8} - 10^{-5} M_\odot \text{ yr}^{-1}$ to spherical winds of speed $w \approx 5 - 15 \text{ km s}^{-1}$ (ref. 11). If the Crab red giant moved subsonically through the hot interstellar medium with speed V , its wind would have terminated behind a strong radiative shock at a distance $r = (\dot{M}w/4\pi p)^{1/2}$ from the star, where $p = \rho s^2 \approx 10^{-12} \text{ dyn cm}^{-2}$, and ρ and s are the mass density and isothermal sound speed of the ambient medium, respectively. Upstream of the star, the shocked wind would be subject to a shear stress $\sim \rho V^2$, assuming a drag coefficient of the order of unity. The shocked wind would be swept downstream with a speed $V^2 w/s^2$ relative to the star. Frictional forces exerted through a turbulent boundary layer would ultimately accelerate the stellar material to the speed of the surrounding interstellar medium; if we estimate the shear stress by ρV^2 , the characteristic acceleration distance is about $(s^2/Vw)r$. The mass profile downstream depends on the detailed flow pattern behind the star, but since the shocked wind is much denser than the interstellar medium, the mass should be concentrated within a hollow cylinder of radius r . A variable discharge would clearly alter the profile; if $\dot{M}$ oscillated, similar structures would be expected on both the brightened limbs of the jet. The observed jet radius, $r \sim 7 \times 10^{17} \text{ cm}$ and an assumed wind speed $w = 5 \text{ km s}^{-1}$ lead to $\dot{M} \sim 2 \times 10^{-7} M_\odot \text{ yr}^{-1}$.

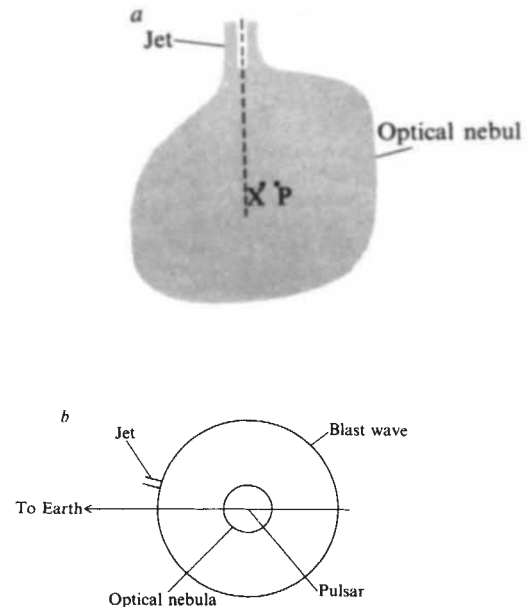


Fig. 1 *a*, Sketch of Crab Nebula and jet projected onto the sky: X, the explosion centre; P, the current pulsar position. *b*, Projection onto plane containing jet and line of sight. If the jet is exterior to a bounding blast wave of ~ 7 pc radius then it might be inclined to the line of sight at an angle $\sim 15^\circ$ as shown. If the blast wave has a radius of ~ 2 pc similar to that of the optical nebula then the jet can be correspondingly shorter.

Let us adopt a stellar velocity $V \approx 30 \text{ km s}^{-1}$, which is compatible with the observed radial velocity of the nebula¹², and is consistent with the Crab's 200 pc distance from the galactic plane. At an inclination of 15° , the observed length of the trail is about 10 pc; its age is $3 \times 10^5 \text{ yr}$, and its mass $0.06 M_\odot$. This estimated age is shorter than the duration of the red giant phase of an $8 M_\odot$ star¹³, but comparable to that for major changes in the evolutionary track in the H-R diagram¹⁴. This age is in good accord with the evaporation lifetime in the hot interstellar medium^{15,16}. For $s \approx 10^2 \text{ km s}^{-1}$, the distance required to accelerate the trail to the speed of the interstellar medium is about 14 pc, so the velocity of the jet may be slightly different from the radial velocity of the Crab Nebula.

The expansion centre of the Crab Nebula's filaments is misaligned by $\sim 10^\circ$ with the jet's symmetry axis. This is unlikely to have been caused by an acceleration of the star just before the supernova. Instead, we attribute it to a difference in flow speed of the interstellar medium of $\sim 5 \text{ km s}^{-1}$ along the trail.

In the absence of ionizing radiation, the gas in the stellar trail will cool to a temperature of $\sim 30 \text{ K}$. It will compress to a density $n \approx 200 \text{ cm}^{-3}$ to achieve pressure equilibrium with the interstellar medium. The thickness of the tube walls, Δr , is given by $\rho \Delta r \approx \dot{M}/2\pi rV$. With $V \approx 30 \text{ km s}^{-1}$, we find $\rho \Delta r \approx 8 \times 10^{-7} \text{ g cm}^{-2}$, and if $n \approx 200 \text{ cm}^{-3}$, $\Delta r \approx 3 \times 10^{15} \text{ cm}$. Turbulence could well broaden the tube thickness. When the trail is subsequently exposed to the UV ionizing flux $\phi \approx 10^9 \text{ cm}^{-2} \text{ s}^{-1}$ from the Crab Nebula, an ionization front will propagate into it with speed $\approx \phi/n \approx 50 \text{ km s}^{-1}$. At this velocity the front will have traversed the tube thickness in the 10^3 yr since the supernova, even allowing for the grazing incidence of the ionizing radiation. The gas should therefore be completely ionized. With a temperature of $\sim 10^4 \text{ K}$, the ionized gas initially has a pressure more than 100 times that of the ambient medium. If it expands freely at the sound speed $\sim 10 \text{ km s}^{-1}$, the tube walls would now be $r \approx 3 \times 10^{16} \text{ cm}$ thick. The mean density should also have decreased by a factor 10 since the supernova.

The present emission measure of the trail, allowing for the inclination and both sides of the tube, is about $10 n^2 \Delta r \approx 40 \text{ cm}^{-6} \text{ pc}$, about 6 times that of the shell⁶, and comparable

to the rough observational estimate in ref. 17. The low column density ensures that there will be no region of low ionization, consistent with the observed high excitation^{1,2,4}.

That part of the trail which has been struck by the supernova ejecta should emit X rays at a rate $L_X \approx 3 \times 10^{34} n_1 \Delta M/M_\odot \text{ erg s}^{-1}$, where ΔM is the mass of the shocked trail, and n_1 is the density in the trail ahead of the blast wave. If $n_1 \approx 20 \text{ cm}^{-3}$ and $\Delta M/M_\odot \approx 0.05$, then $L_X \approx 3 \times 10^{34} \text{ erg s}^{-1}$; gas shocked earlier would have a higher n_1 and L_X . A 0.5–1.5 keV X-ray hot spot has been observed near the location of the optical jet with a luminosity of $\sim 10^{36} \text{ erg s}^{-1}$ (ref. 9). If $\dot{M}$ increased significantly just before the supernova, this could conceivably account for the enhanced soft X rays in the vicinity of the trail.

We therefore conclude that the Crab Nebula jet may be a trail left by the progenitor star. An immediate observational test of this proposal is, in principle, possible. The elemental abundances within the Crab Nebula have been known to be peculiar since the early work of Woltjer¹⁸. Hydrogen is definitely underabundant with respect to helium, with heavier elements either normal or underabundant with respect to total mass¹⁹. No calculated stage of evolution of any normal star can produce the observed abundances and it has been suggested²⁰ for these and other reasons that the Crab progenitor was a runaway star, the second member of a high mass binary system having lost much of its hydrogen envelope in relatively early stages of evolution when the binary environment was important. In any case, retention of some residual hydrogen would have been necessary for evolution to a red (super)giant stage just before explosion. This, in turn, is required in our model to understand the low velocity of the pre-supernova wind (here taken to be 40 km s^{-1}) which is characteristic of the escape velocity from an evolved star with a hydrogen-rich envelope. The wind would have arisen from the outer layers of this envelope. Thus we specifically predict that the hydrogen-to-helium ratio in the jet should be larger than in the nebula as a whole. Due to semi-convection and consequent main sequence mixing the jet may be helium enriched compared with normal matter but, if our model is correct, a relative enhancement of H/He would be expected.

Given the large fraction of high mass stars in relatively close binary systems, and the fact that red giant winds contribute significant mass to the interstellar medium, star trails of the cometary nature we are proposing should be common. If the Galaxy contains 10^6 red giants moving with a speed of 30 km s^{-1} , and the trails survive a million years before evaporating, the average distance to the nearest trail is only $\approx 50 \text{ pc}$. A search for trails in the solar neighbourhood, and attached to nearby giants could be worthwhile. If red giant trails are a widespread interstellar phenomenon, they should provide a powerful diagnostic tool for quantifying the mass loss from red giants, the interstellar velocity field, and our theory of evaporation.

It has been proposed that the jet may be a relic from a supercritical accretion disk, analogous to that in R Aquarii and SS433, from a pre-existing interacting binary system disrupted by the supernova explosion. This is not implausible and would also imply that the jet would be relatively enhanced in hydrogen relative to the nebula. The best test for this proposal is the most obvious one—a deep search for a counterjet which should be seen opposite the jet if this proposal is correct.

We thank F. Coroniti, T. Gull and participants of the Space and Astrophysical Plasmas programme at the Institute for Theoretical Physics for useful discussions, also B. Paczynski for useful discussions. This research was supported by AST80-11752 and the Alfred P. Sloan Foundation (R.B.), NASA NSF-7341 (C.K.), NSF79-23243 (C.F.M.) and PHY27-27084, and NSF AST80-22785 (J.P.O.).

4. Chevalier, R. A., Gull, T. R. *Astrophys. J.* **200**, 399–401 (1975).
5. Chevalier, R. A. in *Supernovae* (ed. Schramm, D.) 53–61 (Reidel, Dordrecht, 1977).
6. Murdin, D. & Clark, D. H. *Nature* **294**, 543–544 (1981).
7. Henry, R. B. C., MacAlpine, G. M. & Kirstiner, R. P. Preprint (Univ. Michigan, 1982).
8. Toor, A., Palmieri, T. M. & Seward, F. D., *Astrophys. J.* **207**, 96–101 (1976).
9. Charles, P. A. & Culhane, J. L. *Astrophys. J. Lett.* **211**, L23–L27 (1977).
10. McKee, C. F. & Ostriker, J. P. *Astrophys. J.* **218**, 148–169 (1977).
11. Zuckerman, B. A. *Rev. Astr. Astrophys.* **18**, 263–288 (1980).
12. Trimble, V. *Astr. J.* **73**, 535–547 (1968).
13. Davidson, K. et al. *Astrophys. J.* **253**, 696–706 (1982).
14. Clayton, D. D. *Principles of Stellar Evolution and Nucleosynthesis*, 491 (McGraw-Hill, New York, 1968).
15. Cowie, L. I. & McKee, C. F. *Astrophys. J.* **211**, 135–146 (1977).
16. Cowie, L. I. & Songaila, A. *Nature* **266**, 501–503 (1977).
17. Davidson, K. & Humphries, R. *Publ. Astr. Soc. Pacif.* **88**, 312–316 (1976).
18. Woltjer, L. *Bull. astr. Inst. Netherlands* **14**, 39–80 (1958).
19. Davidson, K. *Astrophys. J.* **228**, 179–190 (1979).
20. Gott, J. R., Gunn, J. E. & Ostriker, J. P. *Astrophys. J.* **160**, L91–L95 (1970).

White dwarfs and neutron stars in globular cluster X-ray sources

Piet Hut

Institute for Advanced Study, Princeton, New Jersey 08540, USA

Frank Verbunt*

Sterrenkundig Instituut, Utrecht & Sterrenkundig Instituut, Amsterdam, The Netherlands

We predict here that globular clusters contain at least as many cataclysmic variables, which contain white dwarfs, as bright ($L_X > 10^{36} \text{ erg s}^{-1}$) X-ray sources, which contain neutron stars. Globular clusters contain many more white dwarfs than neutron stars, but the capture mechanisms for the formation of binaries are more efficient for neutron stars. We point out here the consequences of the frequent formation of temporarily bound triple systems (resonance scattering). These are: an extra enhancement of neutron star capture with respect to white dwarf capture, a considerable probability of actual collisions between stars, and the presence of cataclysmic variables in the outer regions of globular clusters.

Several cataclysmic variables have been detected optically in globular cluster fields^{1,2}. At least three of these, one dwarf nova³ and two novae, are very good candidates for actual membership of a cluster². This points to a very high true nova rate in globular clusters in comparison with the galactic disk². The two novae and the best dwarf nova candidate are located at distances of 0.6, 0.9 and 10, respectively (in units of the core radius⁴), from the centre.

A preliminary analysis of Einstein IPC data indicated the existence of moderately bright ($L_X = 10^{32} - 10^{34} \text{ erg s}^{-1}$) X-ray sources in globular clusters⁵. This X-ray luminosity is typical for dwarf novae in outburst^{6,7}. Recently a number of moderately bright X-ray sources have been detected with the Einstein IPC detector, typically located at several core radii from the globular cluster centre (J. E. Grindlay and P. Hertz, personal communication).

The presence in globular clusters of binaries containing neutron stars has been explained by two types of capture process. We note that both processes are also possible for white dwarfs. First, a single star encountering a compact star can capture it by dissipating orbital energy. The formation rate of binaries for this process is^{8,9}

$$\mathcal{R}_{\text{cap}} = \frac{N}{10^2 \text{ pc}^{-3}} \frac{n}{10^4 \text{ pc}^{-3}} \frac{M + m}{M_\odot} \frac{x R_*}{R_\odot} \frac{10 \text{ km s}^{-1}}{v_{\text{in}}} \times \\ \times 6 \times 10^{-11} \text{ yr}^{-1} \text{ pc}^{-3} \quad (1)$$

Received 26 July; accepted 25 November 1982.

1. Gull, T. R. & Fesen, R. A. *Astrophys. J. Lett.* **260**, L75–L78 (1982).
2. Fesen, R. A. & Gull, T. R. *Bull. Am. astr. Soc.* **14**, 637 (1982).
3. van den Bergh, S. *Astrophys. J. Lett.* **160**, L27 (1970).

* Present address: Institute of Astronomy, Cambridge CB3 0HA, UK.

where N and n are the number of compact stars and target stars per unit volume and $x \approx 1$ for direct collisions⁸ and $x \approx 3$ for dissipation of the tidal bulge⁹.

Equation (1) gives a higher rate for capture by main sequence stars than by red giants, when proper weighting is performed for numbers, radius and lifetime in different phases of expansion of red giants (R. E. Taam, personal communication).

The second process is the encounter between a single compact star and a binary^{10,11}. (The frequency of the simultaneous encounter of three independently moving stars¹² is negligibly small¹³.) The incoming star can eject one of the binary members immediately on arrival (direct exchange), or a temporarily bound three-body system can be formed, with a typical lifetime of a few hundred original binary periods (P.H., in preparation) (resonance exchange). For hard binaries such as we are interested in, the second process dominates (P.H., in preparation). In the following we will therefore neglect direct exchange scattering. As shown by detailed numerical calculations the rate of resonance scattering between binaries and single stars for equal masses of the three stars is roughly¹¹ (see refs 14–17)

$$\mathcal{R}_{\text{res}} = \frac{N}{10^2 \text{ pc}^{-3}} \frac{n_{\text{bin}}}{10^2 \text{ pc}^{-3}} \frac{m}{M_{\odot}} \frac{a}{1 \text{ AU}} \frac{10 \text{ km s}^{-1}}{v_{\text{in}}} \times \\ \times 5 \times 10^{-10} \text{ yr}^{-1} \text{ pc}^{-3} \quad (2)$$

Here a is the semimajor axis of the binary, N and n_{bin} are the space densities of single stars and binaries, respectively, and m is the mass of each of the three stars. This expression is an approximation that is valid for moderately hard binaries, where the internal orbital velocity exceeds the relative velocity at infinity v_{in} by at least a factor of two.

For white dwarfs equation (2) is appropriate, since their mass is comparable to that of the normal stars in a cluster ($\approx 0.5 M_{\odot}$). For neutron stars, with mass $m_{\text{ns}} \approx 3 m$, say, the rate is roughly twice as high for the same v_{in} (P.H., in preparation).

The fraction of stars that are contained in binaries is less in globular clusters than in the galactic disk^{18,19}. However, even if only 1% of all stars in the cores of globular clusters are in binaries with $a \approx 0.1$ – 1 AU , the formation rate of bright X-ray binaries through exchange scattering can still be high enough to compete with two-body mechanisms, as follows from equation (1) and (2) with $x \approx 3$ and $\langle a \rangle = 0.3$.

We will now estimate the number of binaries formed by capture of a white dwarf instead of a neutron star, through the same mechanisms. For two-body processes we find from equation (1)

$$\frac{\mathcal{R}_{\text{cap,wd}}}{\mathcal{R}_{\text{cap,ns}}} = \frac{M_{\text{wd}} + m}{M_{\text{ns}} + m} \frac{v_{\text{in,ns}}}{v_{\text{in,wd}}} \frac{N_{\text{wd}}}{N_{\text{ns}}} \quad (3)$$

where index wd is for white dwarfs and ns for neutron stars.

For the ratio of resonance scattering events that neutron stars and white dwarfs undergo we write

$$\frac{\mathcal{R}_{\text{res,wd}}}{\mathcal{R}_{\text{res,ns}}} = \frac{1}{2} \frac{v_{\text{in,ns}}}{v_{\text{in,wd}}} \frac{N_{\text{wd}}}{N_{\text{ns}}} \quad (4)$$

where we have used equation (2), taking account of the larger neutron star mass by inserting a factor $\frac{1}{2}$ (see the discussion following equation (2)). Thus, equation (4) turns out to be very similar to equation (3).

Equations (3) and (4) show that the number of neutron stars captured is enhanced relative to that of white dwarfs for both two- and three-body processes, due to (1) the larger mass of the neutron star, (2) their lower velocities (in kinetic equilibrium $\frac{1}{2} M_{\text{ns}} \langle v_{\text{ns}}^2 \rangle = \frac{1}{2} M_{\text{wd}} \langle v_{\text{wd}}^2 \rangle$), (3) their being more concentrated to the dense cluster core than the white dwarfs, and (4) their being older and hence available for longer than were white dwarfs.

Thus, the ratio between two-body formation of binaries for white dwarfs and for neutron stars is approximately the same as the ratio between resonance scattering for white dwarfs and

for neutron stars. Not all resonance scattering events lead to exchange. In the equal mass case, roughly two-thirds of the resonance scattering events result in exchange (P.H., in preparation) as might be expected on statistical grounds. When the incoming star is more massive exchange is even more likely (P.H., in preparation). A neutron star will therefore more often remain in a newly formed binary than will a white dwarf. For very low velocity, however, the relative exchange probability drops because hierarchical resonance scattering which has no exchange, becomes important, (ref. 15 and P.H. and F.V., in preparation).

A further enhancement of neutron star binary formation is due to the fact that binaries are heavier than single stars and will be more concentrated to the central region of the core, in which most of the neutron stars lie. Finally, neutron star binaries will, on formation, have less recoil velocity than white dwarf binaries and have a smaller chance of escaping from the cluster.

The number of white dwarfs per neutron star formed in the history of the globular cluster can be calculated for an initial mass function $dN/dm \propto m^{-x}$ as

$$\left(\frac{m_{\text{min}}}{m_{\text{max}}} \right)^{1-x} \approx \left(\frac{0.8 M_{\odot}}{8 M_{\odot}} \right)^{1-x} = 10^{x-1} \quad (5)$$

where m_{min} is the minimum mass for which globular cluster stars have completely evolved and m_{max} is the maximum mass resulting in the formation of white dwarfs, and above which neutron stars are formed^{20,21}.

The combination of the 4 (7) factors discussed above gives a capture probability per neutron star ~ 10 (30) times higher than the capture probability per white dwarf for $M_{\text{wd}} = 0.5 M_{\odot}$. (For higher M_{wd} the difference is less pronounced.) Combined with equation (5), this gives

$$\frac{N_{\text{bin,wd}}}{N_{\text{bin,ns}}} \approx \frac{1}{f} 10^{x-2} \quad (6)$$

for two-body capture, and

$$\frac{N_{\text{bin,wd}}}{N_{\text{bin,ns}}} \approx \frac{1}{f} 10^{x-2.5} \quad (6')$$

for three-body capture. Here f is the fraction of neutron stars that remain in the globular cluster following the supernova explosion that made them: f is very small if all supernova events lead to high ($\geq 100 \text{ km s}^{-1}$) velocity neutron stars. If the high velocity neutron stars are born in binaries and single massive stars leave low velocity neutron stars (see ref. 22)—as we conclude is indicated by the velocity distribution²³ of the nearest pulsars— f may be of the order of unity. In that case, for shallow slopes in the mass function ($x \approx 2$), roughly equal numbers of white dwarf binaries and neutron star binaries can be expected in globular clusters, while for steeper slopes white dwarf binaries are more abundant.

The number of X-ray binaries observed in a globular cluster depends not only on the formation rate but also on their lifetime as an X-ray source. This causes the expected abundance of X-ray binaries with a red giant companion to be lower than that of binaries with a main-sequence companion, as red giants transfer mass at a fairly high rate and are soon exhausted (S. Rappaport, G. J. Savonije and R. F. Webbink, in preparation). The mass transfer rate of main-sequence companions is not accurately known²⁴. We estimate the fraction of X-ray binaries that contain red giants to be at most a few per cent. Similarly, the observed number of cataclysmic variables is reduced by the following effect: a certain fraction of cataclysmic variables will on formation contain a white dwarf that is less massive than its companion; this leads to an initial stage of mass transfer on a thermal time scale, and a correspondingly shorter lifetime.

There are three observational effects by which the three-body capture mechanism differs from two-body mechanisms. First,

whereas two-body processes are expected to lead to immediately active X-ray binaries because of the small resulting periastron distances, three-body exchange scattering will result in much larger orbits, on the average only slightly tighter than the original binary orbit. The binary thus formed will in general be detached and will light up as an X-ray source only after a considerable period lapse of time, when the companion starts to fill its Roche lobe. This may come about through its evolution into a red giant or through a shrinking of the orbit as a result of angular momentum loss by, for example, magnetic braking²⁴ or wide encounters with field stars²⁵. For large orbits the time scale for shrinking of the orbit will be longer than the evolutionary time scale of the companion, and hence the three-body formation mechanism will enhance the number of red giant companions, and thus the number of very bright X-ray binaries. The fact that the globular clusters in M31 contain bright X-ray sources more luminous than clusters in our own Galaxy²⁶ suggests a larger number of wide initial binaries ($a \approx 0.3$ AU, say) in M31 globular clusters.

Second, during resonance scattering the three stars remain together in erratic orbits for a rather long time, resulting in repeated very close encounters between the stars. Preliminary results show a typical distance of closest approach of $0.02a$ (P.H., in preparation). Obviously, resonance scattering will lead to collisions between stars in many cases, and in some cases even to collisions between all three stars. Collisions between independently moving stars have been suggested to form blue stragglers^{2,3,27}. The formation of blue stragglers through resonance scattering has a larger cross-section (proportional to a rather than to R_*). The frequent occurrence of actual collisions also affects the refuelling process proposed by Shull²⁵, in which the neutron star in an X-ray binary replaces its emptied low-mass companion by a more massive new one. Because the low-mass companion fills its Roche lobe, the orbital separation is very small ($a \approx R_\odot$) in an active X-ray source and the orbital velocity is several hundred km s^{-1} . For such small a values, resonance scattering almost immediately leads to an actual collision with relative velocity of the order of the orbital velocity, and the incoming star is disrupted rather than captured.

The third observational effect of three-body scattering is the wider spread of white dwarf binaries through the globular cluster as compared with neutron star binaries, due to their larger recoil velocities after resonance scattering. Note that both the optical and the Einstein IPC data suggest the existence of cataclysmic variables in the outer regions of globular clusters. The location of a cataclysmic variable in M5 at 10 core radii from the centre is strong evidence for the occurrence of exchange scattering. To improve the statistics we propose a systematic search for novae in the globular cluster system of M87 for which we predict a nova rate of several per year (P.H. and F.V., in preparation).

The report originated at the CECAM workshop on late stages of close binary evolution. We thank the participants of this workshop, especially P. Eggleton, W. H. G. Lewin, R. E. Taam, E. P. J. van den Heuvel and H. van der Woerd, for several stimulating discussions. We thank J. E. Grindlay and P. Hertz for permission to quote their results before publication. P.H. acknowledges support by the NSF through grant PHY79-19884.

Received 25 August; accepted 3 December 1982.

1. Sawyer, H. B. *Publ. David Dunlap Observatory* Vol. 2 No. 2 (1955).
2. Webbink, R. F. *IAU Symp.* No. 88, 561 (1980).
3. Margon, B., Downes, R. A. & Gunn, J. E. *Astrophys. J.* **247**, L89 (1981).
4. Peterson, C. J. & King, I. R. *Astr. J.* **80**, 427 (1975).
5. Grindlay, J. E. in *X-ray Astronomy with the EINSTEIN Satellite*, 79 (ed. Giacconi, R.) (Reidel, Dordrecht, 1980).
6. Cordova, F. A., Mason, K. O. & Nelson, J. E. *Astrophys. J.* **245**, 609 (1981).
7. Cordova, F. A. & Mason, K. O. in *Accretion-driven Stellar X-ray Sources* (eds Lewin, W. H. G. & van den Heuvel, E. P. J.) (Cambridge University Press, in the press).
8. Sutantyo, W. *Astr. Astrophys.* **44**, 227 (1975).
9. Fabian, A. C., Pringle, J. E. & Rees, M. J. *Mon. Not. R. astr. Soc.* **172**, 15p (1975).
10. Hills, J. G. *Mon. Not. R. astr. Soc.* **175**, 1p (1976).
11. Hut, P. & Bahcall, J. N. *Astrophys. J.* (in the press).
12. Clark, G. W. *Astrophys. J.* **199**, L143 (1975).

13. Heggie, D. C. in *Globular Clusters* (ed. Hanes, D. & Madore, B.) 281 (Cambridge University Press, 1980).
14. Hills, J. G. *Astr. J.* **80**, 809 (1975).
15. Heggie, D. C. *Mon. Not. R. astr. Soc.* **173**, 729 (1975).
16. Monaghan, J. J. *Mon. Not. R. astr. Soc.* **176**, 63; **177**, 583 (1976); **179**, 31 (1977).
17. Nash, P. E. & Monaghan, J. J. *Mon. Not. R. astr. Soc.* **184**, 119 (1978).
18. Gunn, J. E. & Griffin, R. F. *Astr. J.* **84**, 752 (1979).
19. Gunn, J. E. in *Globular Clusters* (eds Hanes, D. & Madore, B.) 271 (Cambridge University Press, 1980).
20. Sugimoto, D. & Nomoto, K. *Space Sci. Rev.* **25**, 155 (1980).
21. Romanishin, W. & Angel, J. R. P. *Astrophys. J.* **235**, 992 (1980).
22. Katz, J. I. *Nature* **253**, 698 (1975).
23. Manchester, R. N. & Taylor, J. H. *Astr. J.* **86**, 1953 (1981).
24. Verbunt, F. & Zwaan, C. *Astr. Astrophys.* **100**, L7 (1981).
25. Shull, J. M. *Astrophys. J.* **231**, 534 (1979).
26. Van Speybroeck, L. & Bechtold, J. in *X-ray Astronomy with the EINSTEIN Satellite* (ed. Giacconi, R.) 153 (Reidel, Dordrecht, 1980).
27. Meyer, F. & Meyer-Hofmeister, E. in *Close Binary Stars: Observations and IAU Symp.* No. 88, 145 (1980).

Solar global velocity oscillations and active region rotation

C. J. Durrant & E. H. Schröter

Kiepenheuer-Institut für Sonnenphysik, 7800 Freiburg, FRG

Claverie *et al.*¹ have recently reported the detection of an approximately 13-day periodicity in measurements of the mean Doppler velocity shift of the integrated solar disk. They attribute this signal to the presence of a rapidly rotating core. Here we draw attention to the close correlation between this periodicity and the rotation of solar active regions across the disk. The presence of an active region should modulate the measured velocities due to the temperature sensitivity of the line measured. It seems likely that this is the origin of the signal observed by Claverie *et al.*

The observations of Claverie *et al.*¹ best demonstrate a periodicity in the line-of-sight velocity of the integrated solar disk during the interval 28 June to 22 August 1981 (see Fig. 1a). They find a period of 13.15 days (synodic) and an amplitude of 6.6 m s^{-1} . This period is quite close to being one half of the synodic rotation period of the Sun (27.27 days at $\sim 16^\circ$ lat.). Moreover, it is well known that centres of activity tend to develop at longitudes 180° apart, forming so-called active longitudes². So the passage of activity centres produces quite naturally a periodicity of the correct magnitude. Such a 12-day periodicity has been reported by Knight *et al.*³ in daily sunspot numbers. But sunspots cover too small an area to be suspected of modulating the integrated velocity signal. More likely candidates are the extended plage regions surrounding spot groups. Figure 1c shows the daily degree of activity during this interval, as measured by the calcium index published by the World Data Centre⁴; this is an estimate of the total projected plage area weighted according to brightness on an arbitrary scale. Sine-curve fitting revealed a period of 12.9 days. Note that this is not exactly one-half the solar rotation period because most active regions do not last several rotations. The eruption of new active regions in the centres leads to a drift of their centre of gravity. The centres on opposite sides of the Sun do not drift exactly together so gradual changes of period-phase become apparent in extended records. However, despite the similarity of period, the phase of calcium index lags some 76° behind the velocity signal—the plages are correlated best not when they are at the centre of the disk, but somewhere between the centre and the limb. Under these circumstances it seems unlikely that the modified flow fields in active regions themselves could influence the integrated velocity signal; for one thing, they would be symmetric about the central meridian and would cause the velocity signal to have a turning point there. Instead the signal has opposite sign before and after the calcium index maximum which coincides approximately with central meridian passage. This suggested that the line-of-sight component of the solar rotational velocity was involved.

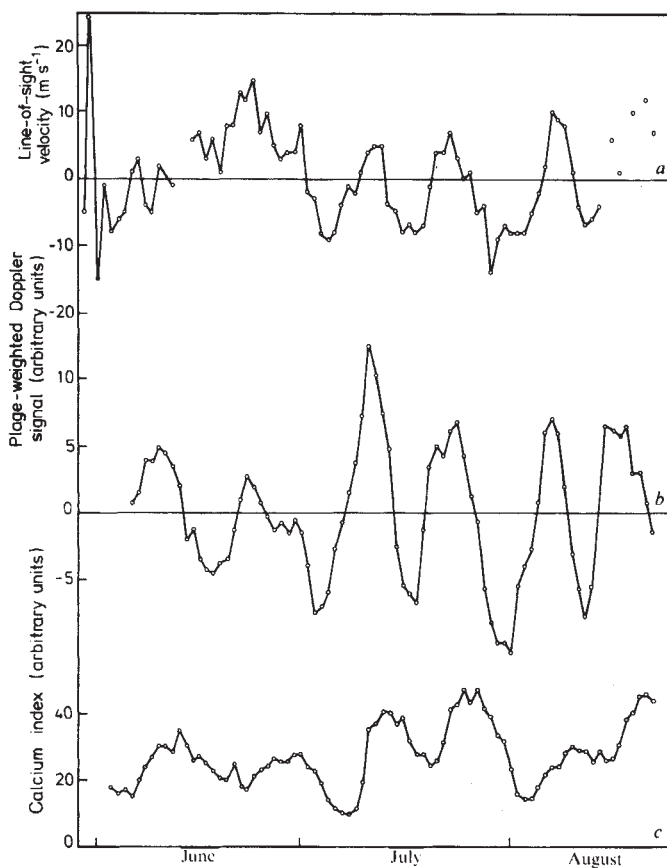


Fig. 1 29 May–22 August 1981: *a*, line-of-sight velocity from ref. 1; *b*, model signal calculated from unbalanced plage rotational velocity component along the line of sight; *c*, daily calcium (activity) index.

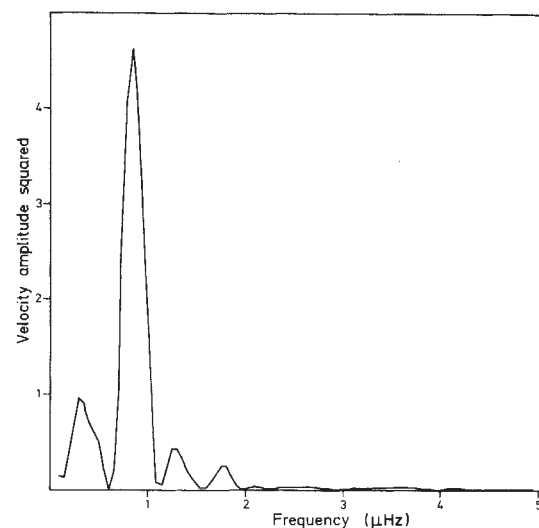


Fig. 2 An amplitude squared spectrum of sine-wave fits to the 28 June–22 August 1981 data.

The measurements are performed by measuring the Doppler shift of the potassium resonance line integrated over the disk. If the strength of this line varies with the degree of activity, the contribution of an active region to the integrated profile would differ from that of a quiet region. If there were to be an active region near one limb and a quiet region at the corresponding opposite limb their contributions to the intensity-weighted Doppler signal would not exactly cancel out. The net imbalance would vanish if the active region lies at the limb, due to foreshortening, and at the centre of the disk, due to the rotational velocity being tangential to the line of sight.

We modelled this effect roughly by constructing a plage-weighted Doppler velocity. For each day, the plage data⁴ were used to calculate a mean line-of-sight velocity by weighting the rotational velocity component of each plage by its projected area and estimated brightness

$$V_p = \alpha \sum_{\text{plages}} V_{\text{rot}}(\text{equator}) \times \cos \theta \sin \phi (A \cos \theta \cos \phi) I \quad (1)$$

where ϕ is the angular distance from the central meridian, θ the latitude, A the area corrected for foreshortening, and I is the visual estimate of the relative plage brightness in arbitrary units⁴. α is a factor that calibrates our model, as we have not considered how the profile changes in active regions nor its influence on the measured Doppler shift; we simply assume, due to the smallness of the effect, a linear behaviour. Comparison of the calculated and observed velocity amplitudes indicates that for a typical plage ($I = 3.0$), $\alpha \sim 0.5$ – 1.0 . This factor is not included in the values shown in Figs 1 and 3, which

are therefore in arbitrary units. The small angle of inclination of the solar rotation axis and limb darkening are neglected. Nevertheless, the model signal (Fig. 1*b*) and the measured velocity signal are remarkably similar. Analysis shows the model to have a period of 13.0 days and a phase shift relative to the measured signal of 10° on 28 June. This shift is less than 1 day. Figure 2 shows a frequency spectrum (the square of the amplitude of least-square sine-wave fits) corresponding to Fig. 3 of ref. 1. The two spectra are very similar. The only dominant peak is at ~ 13 days and its width is consistent with the intrinsic resolution of the 48-day data string. During the interval that the observed signal showed a markedly periodic behaviour, the model signal was equally clear and of large amplitude. Before 28 June, the model signal continues but with much reduced amplitude, while the observed signal is confused.

Comparison of the curves shows that the observed velocity maximum occurs when the plages move across the eastern half of the disk. This implies that the west side (rotational velocity away from the observer) then dominates the integrated profile, and that the potassium line weakens in active regions. As potassium has the extremely low ionization potential of 4.34 eV, the resonance line absorption will occur mostly in cool regions of the temperature minimum and upper photosphere. It will also be extremely sensitive to temperature changes as far as the populations are controlled by collisions and not photoionization⁵. The increased temperature of the photospheric line-forming regions beneath chromospheric plages should indeed weaken the line. To our knowledge this has not been studied either theoretically or observationally. Preliminary results of measurements of the line profile made at the Locarno Observatory by P. N. Brandt indicate that the line does indeed weaken substantially in plages, but the observations do not yet suffice for a quantitative estimate of the effect of the weakening on a velocity measurement like that of Isaak and his collaborators. The value of the factor α derived empirically above is large, so we regard a detailed analysis of the sensitivity of the measurement procedure to be crucial.

As we are unable to provide here a quantitative model we looked also at the plage data for 1977. In this year the degree of activity was very low. The calcium index shown in Fig. 3*b* shows a mean level some 30% of that of 1981. Nevertheless, a strongly varying model velocity signal is still found (Fig. 3*a*) as it is not the total number of active regions on the disk which is important, but the imbalance between the east and west

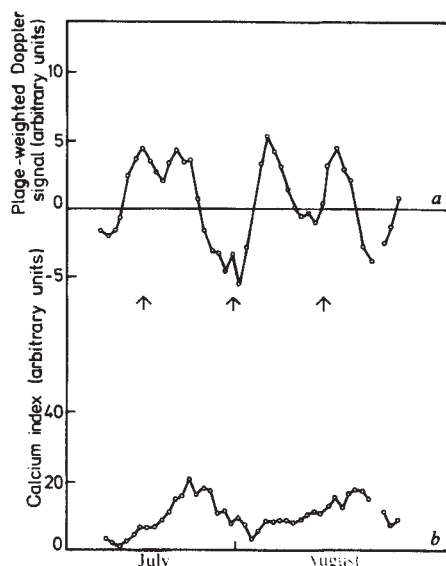


Fig. 3 12 July–24 August 1977: *a*, model signal as Fig. 1; *b*, daily calcium index. The arrows mark the line-of-sight velocity minima reported by Claverie *et al.*¹.

limbs. This imbalance persists even in quiet periods of the cycle; preferred longitudes also occur but are not so long-lived. The signal in Fig. 3*a* is not really periodic, as the appearance of the isolated strong plage regions is more variable. The amplitude of the signal is comparable with that during June 1981, but less than that during July–August.

These results are not easily compared with the data of Claverie *et al.* since they give only the period and phase of a sine-wave fit to their observations. The positions of the observed minima are denoted by arrows in Fig. 3*a*. The reported amplitude is the same as that found in 1981. Whether the apparent phase difference of ~ 2 days is significant cannot be judged without knowledge of the confidence limits of the observational fit.

Finally, a similar analysis of the August 1978 data revealed a minimum in the model velocity signal between 3 and 4 August in agreement with the minimum on 3 August found by Claverie *et al.* We find therefore a strong indication that the 13-day line-of-sight velocity signal is related to the passage across the disk of preferred longitudes of solar activity. If this is indeed the case there is then no evidence for a physical velocity fluctuation with a 12–13-day period. The preferred longitudes, marked by active regions erupting 180° apart in longitude, rotate with a 26-day period, only slightly shorter than that of the surface plasma. The forward drift is also evidenced by the observation that new active regions tend to develop in the preceding part of existing regions. The faster rotation might well reflect that of the deeper layers in which the field is anchored, the 4% difference being well within the limits envisaged by conventional differential rotation. The occurrence of such preferred longitudes still requires to be explained, but this forms part of the general problem of solar dynamo theory for which considerations of the internal rotation and magnetic field distribution of the Sun are of crucial importance. But we would claim that the observations of Claverie *et al.* provide no new direct information regarding these properties.

Received 31 August; accepted 13 December 1982.

1. Claverie, A. *et al.* *Nature* **299**, 704–707 (1982).
2. Dodson, H. W. & Hedeman, E. R. *IAU Symp.* No. 35, 56–63 (1968).
3. Knight, J. W., Schatten, K. H., & Sturrock, P. A. *Astrophys. J. Lett.* **227**, L153–L156 (1979).
4. *Solar-Geophysical Data National Geophysical and Solar-Terrestrial Data Center*, Boulder, Colorado, Year?
5. De la Reza, R. & Müller, E. A. *Sol. Phys.* **43**, 15–32 (1975).

Crystallography of hexagonal germanium

J. R. Parsons & C. W. Hoelke

Materials Science Branch, Atomic Energy of Canada Limited, Chalk River Nuclear Laboratories, Chalk River, Ontario, Canada KOJ 1J0

When a thin foil of amorphous germanium is crystallized by a pulsed laser beam¹, an electron beam or by vacuum annealing², the resulting polycrystalline matrix has the diamond cubic transmission electron diffraction pattern shown in Fig. 1. However, we have recently obtained crystal lattice images from some of the crystals in this polycrystalline matrix which had interplanar angles that were not characteristic of a diamond cubic structure. We now resolve this problem by demonstrating that crystals in the polycrystalline matrix had either a hexagonal or a diamond cubic crystal structure.

Amorphous germanium foils 23 nm thick were formed by vapour deposition onto rocksalt or glass microscope slides at room temperature. Crystallization was accomplished either by a 700 K vacuum anneal while the amorphous foil was still attached to its rocksalt substrate or by electron beam heating in an Elmiskop 101 electron microscope. In the latter case, the amorphous foils so deposited were transferred to 100-mesh electron microscope grids. Then, while simultaneously observing these grid mounted foils, the microscope's condenser aperture was removed and the beam current increased until crystals appeared in the hot central portion of the area exposed to the electron beam. Growth of the crystals was interrupted by replacing the moveable aperture.

Crystal lattice images were recorded once thermal equilibrium was re-established and the specimen had stopped drifting. Lattice images were obtained using 100-kV electrons, axial illumination, single crystal pointed filaments, an objective lens defocus of ~ 300 nm and an electron optical magnification of $\times 445,000$. The objective aperture used was large enough to pass the transmitted beam and the entire first ring of the transmission electron diffraction pattern shown in Fig. 1. The crystal lattice image shown in Fig. 2 was selected as an example of the non-cubic structure of a crystal in the crystallized foil. In Fig. 2 the four sets of resolved planes intersect with interplanar angles of 56° , 62° , 62° and 86° rather than the cubic

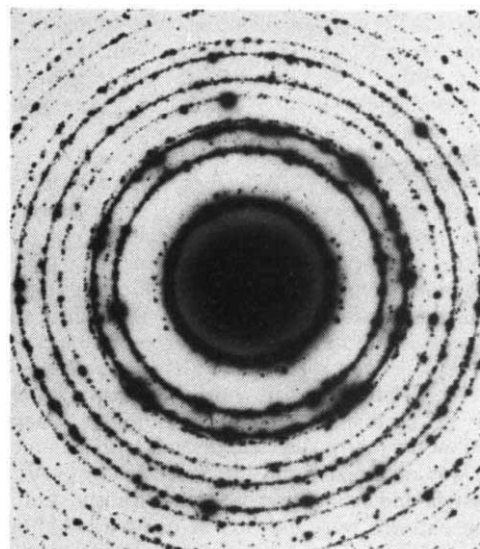


Fig. 1 Transmission electron diffraction pattern of polycrystalline germanium crystallized from its amorphous matrix.

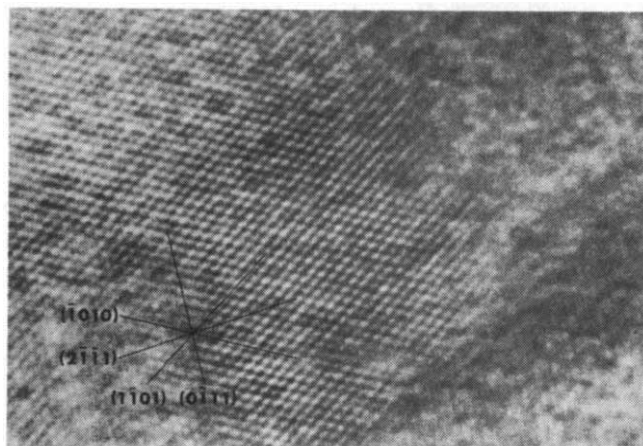


Fig. 2 Crystal lattice image of a $[12\bar{1}3]$ oriented hexagonal germanium crystal.

angles of 55° , 55° , 70° and 90° . Figure 3 is a selected area electron diffraction pattern from a much larger crystal with the same non-cubic structure and orientation.

A trial and error method was used to deduce the structure of these non-cubic crystals. The best fit was achieved when they were given a hexagonal unit cell with dimensions

$$a = 0.396 \pm 0.002 \text{ nm};$$

$$c = 0.980 \pm 0.028 \text{ nm} \quad \text{and} \quad c/a = 2.47$$

With this c/a ratio the crystal lattice image in Fig. 2 and the electron diffraction pattern in Fig. 3 are unique to a $[12\bar{1}3]$ orientation of this proposed hexagonal germanium lattice. In the lattice image a set of (1010) prism planes with a spacing of 0.343 nm is crossed at 62° by a set of 0.324-nm spaced (0111) pyramidal planes which are 56° away from a second set of 0.324-nm spaced (1101) pyramidal planes and 86° away from a set of 0.194-nm spaced (2111) planes. Interplanar angles occurring on lattice images from 30 crystals with this hexagonal structure and orientation have been measured. The average angle between the resolved prism planes and the two sets of resolved pyramidal planes was $61.91 \pm 0.66^\circ$ and the average interplanar angle between the resolved pyramidal planes was $56.10 \pm 0.59^\circ$. The frequency with which crystals having this $[12\bar{1}3]$ hexagonal orientation appeared in lattice images was dependent on the mode of crystallization. For foils crystallized on their rock salt substrate 84 crystal lattice images were examined and only three had interplanar angles characteristic of hexagonal germanium; all the rest had cubic interplanar angles. However, in foils crystallized by electron beam heating, crystals with a $[12\bar{1}3]$ hexagonal orientation appeared as frequently as did crystals with a $[110]$ cubic orientation.

Because this c/a ratio of 2.47 is very close to the 2.45 value for which the corresponding lattice is cubic, transmission electron diffraction patterns from some of the larger crystals in the polycrystalline matrix could be indexed as either hexagonal or diamond cubic orientations of the germanium lattice. For example, crystals with a $[01\bar{1}1]$ or $[01\bar{1}2]$ hexagonal orientation were respectively indistinguishable from crystals with a $[110]$ or $[211]$ diamond cubic orientation. Unambiguous evidence that the matrix also had crystals with a diamond cubic structure was obtained from lattice images of $[110]$ oriented cubic crystals and from crystals whose transmission electron diffraction patterns were indexed as being a $[111]$ or a $[100]$ orientation of the diamond cubic lattice. The latter orientation had missing (200) and (420) reflections characteristic of a diamond cubic structure.

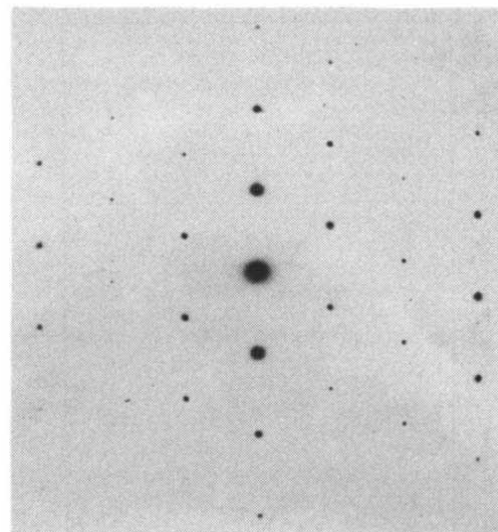


Fig. 3 Transmission electron diffraction pattern from a large crystal of $[12\bar{1}3]$ oriented hexagonal germanium.

We believe that this is the first experimental demonstration that micrometre-sized crystals of germanium can be grown with a hexagonal crystal structure. The fact that crystals with two different crystal structures can coexist in the same polycrystalline matrix at standard temperature and pressure is quite remarkable and suggests that the Gibbs free energy difference between the two germanium crystal structures must be very small. This possibility is consistent with recent calculations of the static structural properties of silicon which includes hexagonal silicon as one of the minimum structural energies³. Also, experimental evidence exists for a hexagonal phase of silicon as precipitates in ion implanted silicon⁴.

Received 6 September; accepted 1 December 1982.

1. Andrew, R. & Lovato, M. *J. appl. Phys.* **50**, 1142–1144 (1979).
2. Parsons, J. R. & Balluffi, R. W. *J. phys. Chem. Solids* **25**, 263–272 (1964).
3. Yin, M. T. & Cohen, M. L. *Phys. Rev. Lett.* **45**, 1004–1007 (1980).
4. Tan, T. Y., Foll, H. & Hu, S. M. *Phil. Mag.* **44**, 127–140 (1981).

Viscosity structure of a layered convecting mantle

Geoffrey F. Davies

Department of Earth and Planetary Sciences and
McDonnell Center for Space Sciences, Washington University,
St Louis, Missouri 63130, USA

Since the formulation of plate tectonics¹, there has been vigorous debate among geophysicists and isotope geochemists concerning the depth extent of mantle flow associated with lithospheric plate motions^{2–9}. Some believed that the flow is confined to the upper mantle^{2,3,6}, others that the flow may penetrate deeper or throughout the mantle^{4,5,8,9}. Recently, it was suggested^{5,10} that if the lower mantle convected separately, then it would probably be so hot that it would melt extensively in its upper portion. As there is no evidence for such melting, it was concluded that whole-mantle circulation was more probable. Here I present a related but more robust argument which leads to the same conclusion. I demonstrate that in a layered mantle, the lower mantle would probably have a viscosity two to three orders of magnitude less than the upper mantle, and this is probably incompatible with observational constraints.

Schubert and Spohn^{9,10} assumed that the lower mantle viscosity is the same as that of the upper mantle. This is compatible

with postglacial rebound observations¹¹, but the resolution of the observations in the lower mantle is not great, and variations by a factor of 10 or more are possible. Furthermore, these workers did not consider the possibility that the solidus of the lower mantle might be considerably higher than that of the upper mantle.

The argument given here can be qualitatively understood by referring to Fig. 1. Because the rheology of silicates is strongly temperature-dependent, the cooler boundary layer at the top of the lower mantle has a higher viscosity. If this boundary viscosity becomes much greater than that in the interior of the lower mantle, the boundary layer will become effectively static, and therefore unavailable to drive lower mantle convection. Thus only that part of the boundary layer that is not too much cooler than the interior will be available to drive the flow. The essential point in this argument is that the surface boundary layer (the lithosphere) is different from the internal boundary layer just described in that none of it is static, even though it is much cooler than the interior of the upper mantle: the surface lithosphere breaks and subducts, thus making the entire upper thermal boundary layer available to drive mantle flow through its (negative) buoyancy forces. I shall show below that the buoyancy forces generated by the internal boundary layer are smaller than those generated by the surface lithosphere, so that the resistance forces (that is, the viscosity) must be correspondingly smaller to transport a given amount of heat. This argument will now be quantified.

Jaupart and Parsons¹² studied numerical models of this type of convection and found that the upper stiff part of the boundary layer becomes static when its viscosity is 100–1,000 times that in the interior of the fluid. Thereafter, viscosity contrasts in the active part of the fluid do not exceed a factor of ~30. The latter value is also evident in the results of Yuen *et al.*¹³, who studied the instability of this type of boundary layer. The temperature dependence of mantle viscosity, η , can be approximated by¹⁴:

$$\eta = \gamma \exp(E/RT) \quad (1)$$

where γ is a constant, E is an activation energy, R is the gas constant and T the absolute temperature. The constant γ can be evaluated from a reference material, denoted by subscript 'r' and defined explicitly below, which may have a different activation energy:

$$\frac{\eta}{\eta_r} = \exp \left[\frac{E_r}{RT_r} \left(\frac{E}{E_r} \frac{T_r}{T} - 1 \right) \right] \quad (2)$$

An analogous relationship implicitly defines the transition temperature, T_t , which separates the active boundary layer from the static part (Fig. 1), in terms of the interior temperature, T_i , and the viscosity ratio, η_t/η_i , at which this transition is assumed to occur:

$$\frac{\eta_t}{\eta_i} = \exp \left[\frac{E_r}{RT_r} \frac{E}{E_r} \left(\frac{T_r}{T_t} - \frac{T_r}{T_i} \right) \right] \quad (3)$$

The boundary layer theory of high Rayleigh number convection relates convected heat flux, q , viscosity, and the temperature difference between the surface and the interior of the fluid, $(T_i - T_s)^{15,16}$:

$$q = A(T_i - T_s)^{4/3} \eta^{-1/3} \quad (4)$$

where A is nearly independent of the depth of the fluid and is otherwise a constant. We can now assume that the upper mantle is the reference system, thus effectively evaluating both A and γ in terms of the following assumed values: $T_{sr} = 270$ K, $T_r = 1,670$ K, $E_r/RT_r = 30$ (ref. 17). To a sufficient approximation, the surface heat flux, q_r (excluding near-surface radiogenic heat in the continental crust), can be taken to be equal to the heat flux out of the lower mantle, q_l ; only a small part of the surface flux is generated in the upper mantle^{9,10}, and this is compensated

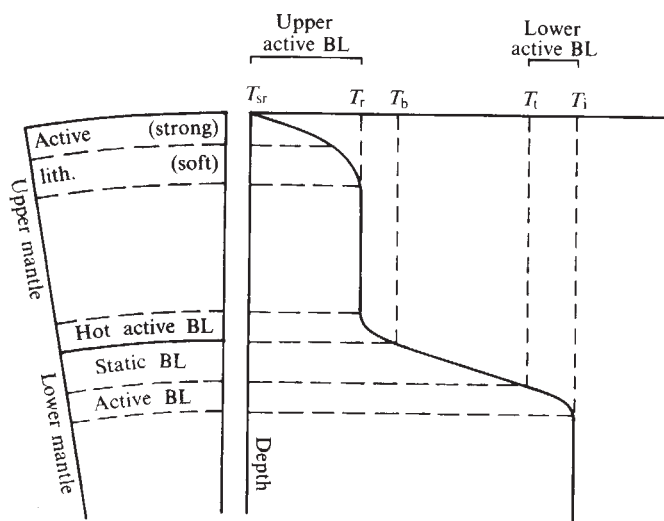


Fig. 1 Schematic thermal and mechanical structure of a mantle with two separately convecting layers. Each layer has an upper boundary layer (BL) which is cooler and stronger than the interior of the layer. The top layer also has a warmer, less viscous boundary layer at its base. Part of the boundary layer at the top of the lower mantle may be so much more viscous than the interior that it is static. Then, only that part with temperatures between a transition temperature, T_t , and the interior temperature, T_i , will be active and available to drive convective flow. This is in contrast to the top boundary layer of the upper mantle (that is, the lithosphere), which is observed to break and subduct even though it is much cooler than the interior of the upper mantle. T_r and T_{sr} are defined in the text. T_b is the temperature at the boundary between the layers.

by the smaller surface area of the lower mantle. Using equation (4), this leads to the condition

$$\left(\frac{T_i - T_t}{T_r - T_{sr}} \right)^{4/3} \left(\frac{\eta_i}{\eta_r} \right)^{-1/3} = \frac{q_l}{q_r} = 1 \quad (5)$$

For given values of η_i/η_r and E/E_r , equations (2), (3) and (5) can be solved for the lower mantle viscosity, η_i , and temperature, T_i , and the transition temperature, T_t . This approach amounts to using the upper mantle to calibrate the lower mantle convection system, and to asking how the temperature would have to be adjusted to deliver the same heat flux if the activation energy or the upper thermal boundary layer were changed. The results of solving equations (2), (3) and (5) are illustrated in Fig. 2. The required internal temperatures are substantially higher and the internal viscosities are substantially lower in all cases. Increasing E/E_r by 50% increases the internal viscosity by only about a factor of three, while the assumed boundary viscosity contrast, measured by η_t/η_i , has a stronger effect. In all cases, the internal viscosity is much lower, by factors of 30–2,000, than in the reference (upper mantle) system. These results are in accord with the qualitative arguments given above.

The activation energy for high-temperature creep in olivine is estimated experimentally to be 111–130 kcal mol⁻¹ (ref. 17), while activation energies for close-packed oxides range from 92 to 152 kcal mol⁻¹ (ref. 18), with MgO (a major chemical component of the mantle) having 110–117 kcal mol⁻¹ (ref. 18). These values suggest that the activation energy of the mantle may increase by 0–20% through the transition zone. Sammis *et al.*¹⁸ use an empirical correlation of activation energy with volume per oxygen ion to estimate that the activation energy increases by less than 10%. Thus the results for $E/E_r = 1.0$ and 1.2 are likely to bound the reasonable possibilities for the lower mantle. If $\eta_t/\eta_i = 1,000$, as suggested earlier, the lower mantle viscosity would be about a factor of 100 less than that of the upper mantle (Fig. 2). If η_t/η_i is as small as 30 (refs 12, 13), the viscosity contrast would be a factor of 1,000.

For $E/E_r = 1.0$, the lower mantle temperature would be 2,000–2,200 K (the adiabatic gradient is ignored here, so this

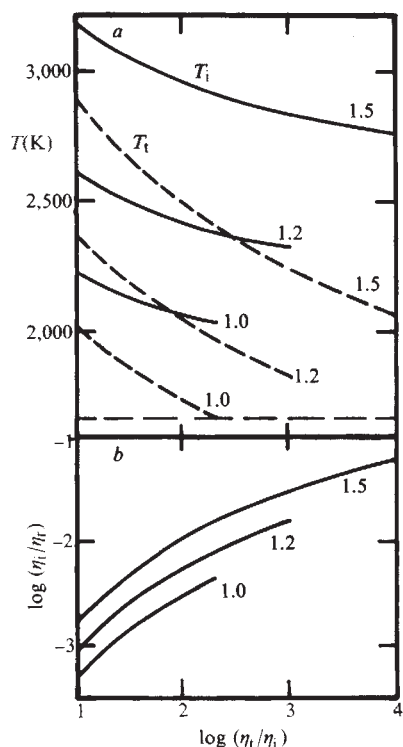


Fig. 2 *a*, Interior temperature, T_i (solid lines) and transition temperature, T_t (broken lines) as functions of the ratio of the viscosity at the transition from active to static boundary layer, η_t , to the interior viscosity, η_i . Curves are labelled with the activation energy ratio E/E_r . *b*, Interior viscosity, η_i , relative to the reference (upper mantle) viscosity, η_r , as a function of the boundary viscosity ratio, η_t/η_i . Curves are labelled with values of E/E_r .

can be interpreted as the temperature near the top of the lower mantle). For $E/E_r = 1.2$, this range increases to 2,300–2,600 K. Especially for the latter case, it is possible that the lower mantle solidus is substantially exceeded^{9,10}, and so the argument used by Schubert and Spohn^{9,10} may still apply. It must be remembered, however, that a higher creep activation energy may be accompanied by higher melting temperatures, as both depend on interatomic bond strengths. Added to the already considerable uncertainty in melting temperatures at high pressures, this suggests that the viscosity predicted by this model may be more amenable to observational testing than is the temperature.

Transition temperatures, T_t , are also shown in Fig. 1, and it can be seen that the temperature difference across the active boundary layer, $(T_i - T_t)$, is less than 500 K for $E/E_r < 1.2$. This is much less than the analogous temperature difference of ~1,400 K across the surface lithosphere, and demonstrates that the buoyancy forces derived from the internal boundary layer are smaller than those derived from the lithosphere.

Recent models of mantle viscosity have incorporated gravity anomalies and secular drift of the rotation pole and rotation rate as constraints in addition to postglacial rebound observations^{19–21}. The preferred models have either uniform viscosities of 10^{22} – 10^{23} P, or an increase of viscosity through the transition zone of at most a factor of 10, from $\sim 10^{22}$ to 10^{23} P. The hypothesis that the lower mantle has a substantially lower viscosity than the upper mantle does not seem to have been tested to date. The fact that the lower mantle viscosity predicted here is two to four orders of magnitude less than in present models based on a recently enlarged data sets suggests that the prediction may be incompatible with the observational constraints. A quantitative test of the low viscosity lower mantle hypothesis may provide a strong test of the layered mantle hypothesis.

This work was supported in part by NSF grant EAR-8212993.

Received 25 October; accepted 8 December 1982.

1. Isacks, B., Oliver, J. & Sykes, L. R. *J. geophys. Res.* **73**, 5855–5899 (1968).
2. Richter, F. M. *Rev. Geophys. Space Phys.* **11**, 233–287 (1973).
3. McKenzie, D. P., Roberts, J. M. & Weiss, N. O. *J. Fluid Mech.* **62**, 465–538 (1974).
4. Davies, G. F. *Geophys. J. R. astr. Soc.* **49**, 459–485 (1977).
5. Peltier, W. R. *Geophys. J. R. astr. Soc.* **46**, 669–706 (1976).
6. Wasserburg G. J. & DePaolo, D. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3594–3598 (1979).
7. O'Nions, R. K., Evenson, N. M. & Hamilton, P. J. *J. geophys. Res.* **84**, 6091–6101 (1979).
8. Davies, G. F. *Nature* **290**, 208–213 (1981).
9. Schubert, G. & Spohn, T. *Geophys. Res. Lett.* **8**, 951–954 (1981).
10. Spohn, T. & Schubert, G. *J. geophys. Res.* **87**, 4682–4696 (1982).
11. Peltier, W. R. in *Physics of the Earth's Interior*, Enrico Fermi int. School Phys., Course 78 (eds Dziewonski, A. & Boschi, E.) 362–431 (Elsevier, Amsterdam, 1980).
12. Jaupart, C. & Parsons, B. *Geophys. J. R. astr. Soc.* (submitted).
13. Yuen, D. A., Peltier, W. R. & Schubert, G. *Geophys. J. R. astr. Soc.* **65**, 171–190 (1981).
14. Weertman, J. *Rev. Geophys. Space Phys.* **8**, 145–168 (1970).
15. Turcotte, D. L. & Schubert, G. *Geodynamics* (Wiley, New York, 1982).
16. Davies, G. F. *J. geophys. Res.* **85**, 2517–2530 (1980).
17. Carter, N. L. *Rev. Geophys. Space Phys.* **14**, 301–360 (1976).
18. Sammis, C. G., Smith, J. C., Schubert, G. & Yuen, D. A. *J. geophys. Res.* **82**, 3747–3761 (1977).
19. Nakiboglu, S. M. & Lambeck, K. *Geophys. J. R. astr. Soc.* **62**, 49–58 (1980).
20. Wu, P. & Peltier, W. R. *Geophys. J. R. astr. Soc.* **70**, 435–485 (1982).
21. Sabadini, R., Yuen, D. A. & Boschi, E. *J. geophys. Res.* **87**, 2885–2903 (1982).

Effects of diagenetic recrystallization on $^{18}\text{O}/^{16}\text{O}$ values of deep-sea sediments

John S. Killingley

Scripps Institution of Oceanography, La Jolla, California 92093, USA

The oxygen isotopic record in deep-sea sediments is a chief source of information on ocean and climate history of the Tertiary period. Traditionally interpretations have focused on the relative importance of temperature change and of polar ice buildup as dominant generating factors^{1–6}. Effects of diagenesis have been largely ignored. Here I show that recrystallization must not be neglected: with reasonable assumptions it can account for much of the familiar Cenozoic trend in the oxygen isotopic composition of foraminifera. Recent studies^{7,8} indicate that there may be considerable recrystallization of Tertiary carbonates with very significant effects on isotopic compositions. Specifically, the work of Baker *et al.*⁷ suggests that there could be total recrystallization in sediments deposited as recently as the early Tertiary.

Our knowledge of the sequence of events during the Cenozoic, as reflected by its isotope stratigraphy, is largely due to the studies of Douglas and Savin¹ and Shackleton and Kennett². Their oxygen isotopic curves for planktonic and benthic foraminifera are shown in Figs 1a and 2a. Here the $^{18}\text{O}/^{16}\text{O}$ ratios are expressed by the usual δ notation as the ‰ difference from a reference standard. A cooling trend, indicated by increasing $\delta^{18}\text{O}$ values, is apparent in the isotopic temperatures of intermediate and high latitudes stretching from the beginning of Tertiary time to the present (Fig. 2a). On this general isotopic trend are superimposed fluctuations, indicative of rather abrupt climatic changes at times. The net increase in benthic $\delta^{18}\text{O}$ values from both the Douglas-Savin and Shackleton-Kennett curves can be interpreted as a decrease in bottom water temperature of about 15 °C from the Palaeocene Epoch to the mid-Miocene. At this time the growth of a substantial Antarctic ice sheet is thought to have increased the $\delta^{18}\text{O}$ values of ocean water by ~1‰ to its present value of ~0‰ (SMOW)³. However, we must now entertain the possibility that much of the oxygen isotope trend is produced by diagenesis.

For a first-order estimate of possible effects of isotopic re-equilibration on the $^{18}\text{O}/^{16}\text{O}$ ratio of biogenic calcite with increasing sediment age, I have constructed a simple model with the following assumptions: (1) recrystallization is a single-step process in a closed system with constant porosity and constant %CaCO₃; (2) the geothermal temperature gradient is

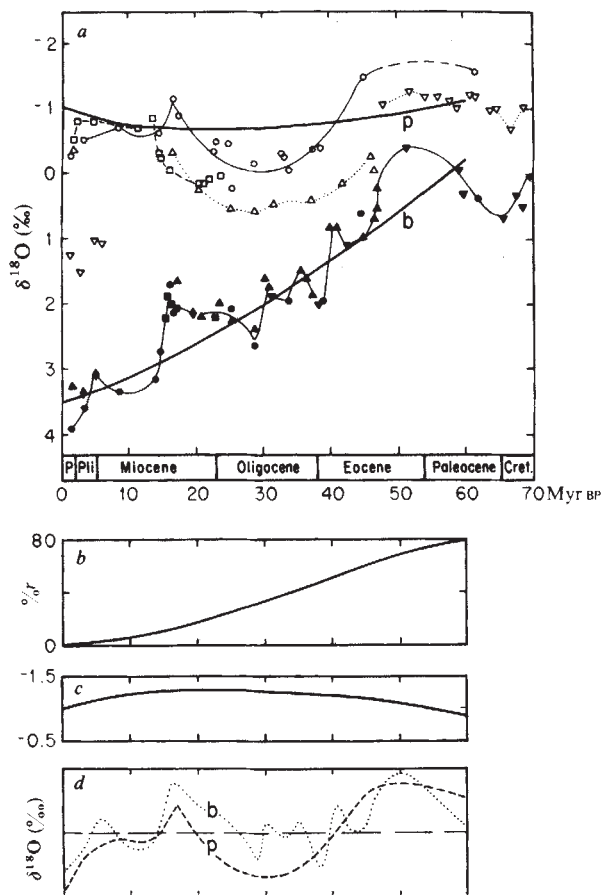


Fig. 1 *a*, Isotopic data for Tertiary planktonic foraminifera (open symbols) and benthic foraminifera (closed symbols) primarily for the North Pacific³. The heavy trend curves are calculated from the diagenetic recrystallization model to predict trends for planktonic (p) and benthic (b) foraminiferal $\delta^{18}\text{O}$ values. (All $\delta^{18}\text{O}$ values are referred to PDB except for pore water $\delta^{18}\text{O}$ values which are referred to SMOW.) *b*, Assumed percentage recrystallization (%*r*) with age. The sigmoid shape of the recrystallization curve from 0 to 80% was chosen to be similar to that which one would obtain by plotting calculated % recrystallization (*r*) values versus depth of Baker *et al.*⁷ for DSDP Core 288. Temperatures used in the calculation (to obtain α_T) were obtained by reference to this recrystallization curve and to the respective temperature/age curve for a given sedimentation rate. *c*, Calculated $\delta^{18}\text{O}$ of pore water as a function of recrystallization. *d*, Observed isotopic curves for benthic (b) and planktonic (p) foraminifera normalized to their respective calculated diagenetic trend curves (horizontal dashed line).

Table 1 Assumed sediment properties for various model runs

No.	Sedimentation rate (cm kyr ⁻¹)	Porosity (%)	CaCO ₃ (%)	δw_i (‰)	$\delta c_i (= \delta p)$ (‰)	δb (‰)
1	1.0	60	90	-1.0	-1.0	3.5
2	0.5	60	90	-1.0	-1.0	3.5
3	1.5	60	90	-1.0	-1.0	3.5
4	1.0	40	90	-1.0	-1.0	3.5
5	1.0	80	90	-1.0	-1.0	3.5
6	1.0	60	10	-1.0	-1.0	3.5
7	1.0	60	80	-1.0	-1.0	3.5
8	1.0	60	90	0.0	-1.0	3.5
9	1.0	60	90	-1.0	-0.5	3.5
10	1.0	60	90	-1.0	-1.5	3.5
11	1.0	60	90	-1.0	-1.0	3.0
12	1.0	60	90	-1.0	-1.0	4.0
13	1.0	60	90	-1.0	2.5	3.5

(δp = planktonic value; δb = benthic value).

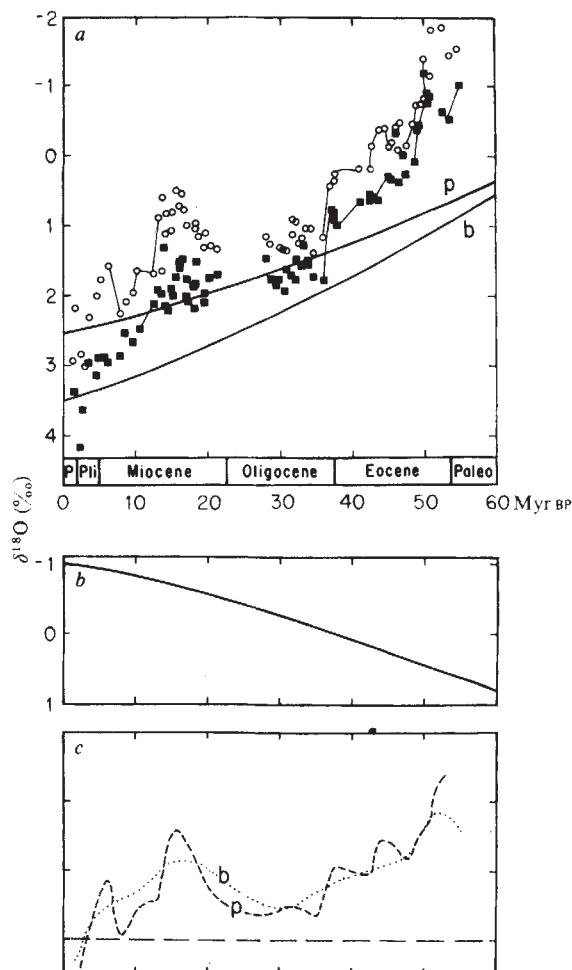


Fig. 2 *a*, Isotopic data for Tertiary planktonic foraminifera (open circles) and benthic foraminifera (closed symbols) from the sub-Antarctic Pacific². The heavy trend curves are calculated from the diagenetic recrystallization model to predict trends for planktonic (p) and benthic (b) foraminiferal $\delta^{18}\text{O}$ values. *b*, Calculated $\delta^{18}\text{O}$ of pore water as a function of recrystallization. *c*, Observed isotopic curves for benthic (b) and planktonic (p) foraminifera normalized to their respective calculated diagenetic trend curves (horizontal dashed line).

equal to 35° km^{-1} , starting with a bottom water temperature of 2°C ; (3) recrystallization is 80% complete at 60 Myr BP (a 'bad', but not the 'worst' case of Baker *et al.*⁷); (4) isotopic equilibration is between carbonate with an initial average value equal to the planktonic value and seawater of a given initial $\delta^{18}\text{O}$ value, where the initial values are invariant through time; (5) recrystallization rates are the same for planktonic and benthic foraminiferal tests.

We have, then, a closed system containing M_c moles of calcite and M_w moles of water with initial $\delta^{18}\text{O}$ values of δc_i and δw_i respectively. It follows, from mass balance considerations, that if *r*% of the calcite recrystallizes in thermodynamic equilibrium with water at a given temperature then δw_e , the $\delta^{18}\text{O}$ value of water in the re-equilibrated system, is given by:

$$\delta w_e = \frac{M_c r [\delta c_i + 10^3(1 - \alpha_T)] + M_w \delta w_i}{M_w + M_c \cdot \alpha_T} \quad (1)$$

where α_T is the fractionation factor between calcite and water at *T* K. This expression is essentially the same as that given by Lawrence².

In the calculations using equation (1) I used a value of either 0‰ (SMOW) or -1‰ (SMOW) for δw_i and various

values of δc_i . PDB values of calcite were converted to SMOW values by means of the expression $\delta^{18}\text{O}_{\text{SMOW}} = 1.03086 \delta^{18}\text{O}_{\text{PDB}} + 30.86$ (ref. 10). α_T values at various temperatures were obtained using the relationship $10^3 \ln \alpha_T = 2.78(10^6 T^{-2}) - 2.89$ (ref. 11).

For each set of calculations I assumed a constant porosity (ϕ), a constant % CaCO_3 , and that the solid phase had a uniform density of 2.7 g cm^{-3} . Thus a value for M_c and M_w could be obtained. Calculations of δw_e were made in 16 steps, assuming that, for each step, 5% of the calcite recrystallized.

The δc_e value of the calcite recrystallized at each step was calculated using the equation

$$\delta c_e = (\delta w_e + 10^3 \ln \alpha_T) 0.97006 - 29.94 \quad (2)$$

[Equation (2) follows from the relationship,

$$\delta^{18}\text{O}_{\text{PDB}} = \delta^{18}\text{O}_{\text{SMOW}} 0.97006 - 29.94 \text{ (ref. 10)}$$

and

$$\alpha_T = \frac{10^3 + \delta c}{10^3 + \delta w}$$

and the approximation $10^3 \ln (1.00X) \approx X$.] Having calculated the δc_e value for 16 steps in each series I then calculated the corresponding recrystallization value at each step starting with an appropriate initial planktonic or benthic $\delta^{18}\text{O}$ value.

Using appropriate ranges of the variables, values were selected (Table 1) to test the model for sensitivity to changes in each of the six main parameters.

Figure 1a shows the isotopic data for the North Pacific^{1,12-14} as represented in ref. 3. Superimposed on the Douglas-Savin plot I have drawn the calculated isotopic recrystallization curves for the 'reference' condition (no. 1 in Table 1) assuming an initial planktonic $\delta^{18}\text{O}$ value of -1% and an initial benthic value of 3.5% (PDB). Figure 1b shows the percentage recrystallization assumed, and the model-derived pore water $\delta^{18}\text{O}$ values are given in Fig. 1c. Also, one of the Douglas-Savin planktonic curves and their benthic curve are redrawn normalized to the respective calculated diagenetic curves (Fig. 1d). Corresponding plots of isotopic data from Shackleton and Kennett² and model-derived curves are shown in Fig. 2.

The fit of the calculated curves with the observed trends, under the more or less reasonable assumptions made, is such that diagenesis can explain the entire Tertiary trend seen. Although this is probably not the case it is important to recognize that diagenesis must be considered as a third major factor besides temperature and ice buildup when discussing Tertiary trends. If the model has general validity then it follows that the convergence of mid-latitude benthic and planktonic signals with age owes much to the effect of diagenetic change in the benthic $\delta^{18}\text{O}$ values.

Although the values of the variables chosen are not unreasonable for deep-sea sediments the choice of 80% recrystallization at 60 Myr is arbitrary. Both faster or slower recrystallization rates could be chosen, based on available information^{7,8}. The fit to the trends in the Douglas-Savin plot would then deteriorate. However, the main thrust of my argument would not be affected, as long as recrystallization rates are not decreased by more than a factor of about three.

Major features of $\delta^{18}\text{O}$ variations are in agreement between different cores (Figs 2d and 3c) and with non-isotopic evidence^{3,15}. This suggests that although re-equilibration with pore water may be occurring, the original isotopic signal is not completely erased because sequential recrystallization is occurring in a closed system or, at least, advection of fluids through the sediment does not have a significant effect.

The degree to which pore water composition may change due to advective and diffusive processes is possibly very variable from site to site and with depth in any given site. Evidence for diffusion and advection in deep-sea cores is contradictory: For example, Friedman and Hardcastle¹⁶ used HDO studies in deep-sea cores spanning the Tertiary and found that water

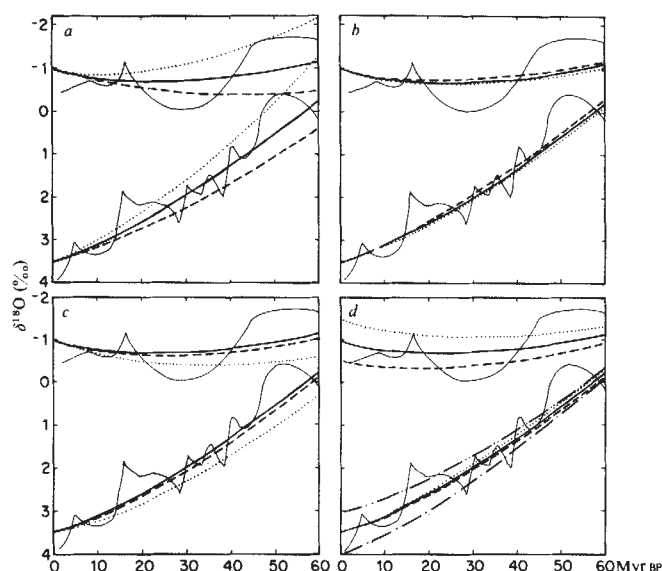


Fig. 3 Sensitivity of diagenesis model to changes in main variables. *a*, Sedimentation rate cm kyr^{-1} --- 0.5; — 1.0; 1.5. *b* Porosity (%) --- 40; — 60; 80. *c* CaCO_3 (%) --- 10; — 90; $\delta^{18}\text{O}$ water, initial: = 0. *d*, $\delta p\%$: -1.5; — -1.0; --- -0.5, with $\delta b = 3.5$; $\delta b\%$: ---, 3.0; — 3.5; --- 4.0, with $\delta p = -1.0$.

leaving the compacting sediment at depth does not mix with overlying pore water but migrates upwards through cracks. They also concluded that self-diffusion of water was an extremely slow process. Recently Sayles and Jenkins¹⁷ have shown that solution advection is apparently occurring through equatorial Pacific sediments at a rate of about 20 cm yr^{-1} . By contrast, other results (ref. 18) indicate little if any flow through a relatively thick ($> 100 \text{ m}$) sediment column.

The calculations (Fig. 3) show that changes in % CaCO_3 between 10 and 90% and in porosity between 40 and 80% do not have marked effects on the model results. The most important effect is the change in temperature/age gradient related to changes in sedimentation rate.

The model for isotopic change presented here is, of course, an oversimplification and there are numerous factors that will affect the postdepositional isotopic record. As Lawrence⁹ has indicated there are potentially large changes in pore water $\delta^{18}\text{O}$ values that can be caused by diagenesis of siliceous materials. He measured the $\delta^{18}\text{O}$ values of pore water in a deep-sea core and noted a decrease of $\sim 2\%$ coincident with the first appearance of a chert horizon in the late Eocene. Effects of such diagenetic changes in pore water $\delta^{18}\text{O}$ would tend to shift both planktonic and benthic $\delta^{18}\text{O}$ signals to even more negative values in the early Tertiary with a resultant increase in the apparent temperature drop across the Tertiary. However, as a further complication, laboratory experiments with pressure solution and hydrothermal recrystallization of carbonates¹⁹ have shown that clay minerals retard the rate of recrystallization.

Diagenesis has occasionally been used by other workers^{9,13,20-26} to warn of possible complications to palaeoconstructions from Tertiary (and older) sediments. Here I have attempted to demonstrate a possible quantitative relationship between the degree of diagenesis and $\delta^{18}\text{O}$ changes. Pre-Miocene stable isotopes provide our main entry into modelling the oceanography and climatology of an ice-free world. We must constrain the effects of diagenesis to be able to proceed with such studies.

I thank W. H. Berger, R. S. Keir and E. Vincent for fruitful discussions. This work was supported by NSF grant OCE80-24610, and ONR contract N00014-80-C-044.

Received 13 October; accepted 6 December 1982.

1. Douglas, R. G. & Savin, S. M. *Init. Rep. DSDP Leg 32*, 509–520 (1975).
2. Shackleton, N. J. & Kennett, J. P. *Init. Rep. DSDP Leg 29*, 743–756 (1975).
3. Savin, S. M. A. *Rev. Earth planet. Sci.* **5**, 319–355 (1977).
4. Matthews, R. K. & Poore, R. Z. *Geology* **8**, 501–504 (1980).
5. Douglas, R. G. & Woodruff, F. C. in *The Oceanic Lithosphere, The Sea Vol. 7* (ed. Emiliani, C.) 1233–1327 (Wiley, New York, 1981).
6. Vincent, E. & Berger, W. H. in *The Oceanic Lithosphere, The Sea Vol. 7* (ed. Emiliani, C.) 1025–1119 (Wiley, New York, 1981).
7. Baker, P. A., Gieskes, J. M. & Elderfield, H. J. *sedim. Petrol.* **52**, 71–82 (1982).
8. Miller, K. G. & Curry, W. B. *Nature* **296**, 347–352 (1982).
9. Lawrence, J. R. *Init. Rep. DSDP Leg 15*, 891–899 (1973).
10. Friedman, I. & O'Neil, J. R. in *Data of Geochemistry, Sixth Edition, III, Title IV, Series V. Paper 440-KK* (US Geological Survey, 1977).
11. O'Neil, J. R., Clayton, R. N. & Mayeda, T. K. *J. chem. Phys.* **51**, 5547–5588 (1969).
12. Douglas, R. G. & Savin, S. M. *Init. Rep. DSDP Leg 6*, 1123–1127 (1971).
13. Douglas, R. G. & Savin, S. M. *Init. Rep. DSDP Leg 17*, 591–605 (1973).
14. Savin, S. M., Douglas, R. G. & Stehli, F. G. *Bull. geol. Soc. Am.* **86**, 1499–1510 (1975).
15. Thompson, G. R., Fields, R. W. & Alt, D. *Geology* **10**, 413–417 (1982).
16. Friedman, I. & Hardcastle, K. *Init. Rep. DSDP Leg 15*, 901–903 (1973).
17. Sayles, F. L. & Jenkins, W. J. *Science* **217**, 245–248 (1982).
18. Crowe, J. & McDuff, R. E. *Earth planet. Sci. Lett.* (in the press).
19. Baker, P. A., Kastner, M., Byerlee, J. D. & Lockner, D. A. *Mar. Geol.* **38**, 185–203 (1980).
20. Milliman, J. D. *Science* **153**, 994–997 (1966).
21. Anderson, T. F. & Schneidermann, N. *Init. Rep. DSDP Leg 15*, 795–803 (1973).
22. Matter, A., Douglas, R. C. & Perch-Nielsen, K. *Init. Rep. DSDP Leg 32*, 891–921 (1975).
23. McKenzie, J., Bernoulli, D. & Garrison, R. E. *Init. Rep. DSDP Leg 42*, 473–477 (1978).
24. Keigwin, L. D. Jr *Earth planet. Sci. Lett.* **45**, 361–382 (1979).
25. Vergnaud-Grazzini, C. *Init. Rep. DSDP Leg 47-2*, 507–511 (1979).
26. Matthews, R. K., Curry, W. B., Lohmann, K. C., Sommer, M. A. & Poore, R. Z. *Nature* **283**, 555–557 (1980).

Isotopic evidence for UK Upper Permian mineralization by bacterial reduction of evaporites

Gill. M. Harwood*

Department of Earth Sciences, The Open University,
Milton Keynes MK7 6AA, UK

Max Laurence Coleman

Isotope Geology Unit, Institute of Geological Sciences,
64 Gray's Inn Road, London WC1X 8NG, UK

The dolomites ($\text{Ca Mg}(\text{CO}_3)_2$) of the Lower Magnesian Limestone (or Cadeby Formation, Upper Permian age) are believed to have been formed by evaporation of seawater and include sulphate also associated with that environment. Mineralization in the formation includes baryte (BaSO_4) and galena (PbS) together with other sulphides and has been ascribed to many different modes of origin. Sulphur isotope analyses have now shown that the sulphates were derived from seawater of Upper Permian age and that sulphides in sites of former anhydrite (CaSO_4) nodules were formed by bacterial reduction of sulphate. This latter point is confirmed by the carbon isotope composition of the carbonates whose oxygen isotope values indicate the possibility of admixture of meteoric water to the system. Other baryte mineralization directly involves Upper Permian seawater or evaporite sulphate, whereas smaller baryte occurrences and epigenetic galena mineralization show no such direct involvement.

Mineralization within the Magnesian Limestone (Cadeby Formation) is both syngenetic and epigenetic with several different mineral associations present¹. Syngenetic (or penecon-temporaneous) mineralization includes replacive and/or void filling galena near the base of the formation and sphalerite-galena-marcasite-baryte-calcite present in the sites of replaced former anhydrite nodules². Epigenetic mineralization comprises void-occluding galena with calcite, dispersed baryte with few partially replacive sulphides, large scale replacive baryte with rare sulphides and fluorite and some cross-cutting baryte veins,

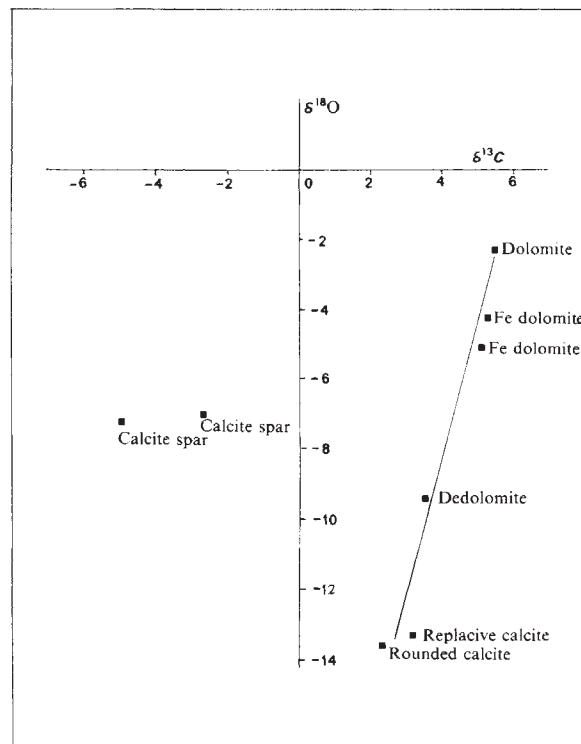


Fig. 1 Carbon-oxygen isotope plot, showing the trend from dolomite to replacive calcite and distinct values of calcite spar.

and galena-wulfenite-asphaltite occurrences at the top of the formation³. Isotope analyses were carried out to try to substantiate the observed petrographic relationships of both sulphides and carbonates in the replaced anhydrite nodules and to determine if a direct relationship existed between evaporite sulphate within and above the formation and the sulphate/sulphide present in the different mineral associations.

Most carbonate, sulphate and sulphide samples were hand scratched from cut slabs. However, microscopic intergrowths of ferroan dolomite and dedolomite (calcite) made hand picking impracticable; the intergrowths were finely crushed and separated by heavy liquid mixtures (ref. 4, appendix B). Carbon and oxygen isotopes were analysed following techniques and corrections reported elsewhere⁵⁻⁸. Carbon and oxygen results have an analytical uncertainty (difference between complete duplicate analyses) of ~0.02‰ and 0.03‰ respectively and are given with reference to PDB. Sulphur isotope measurement techniques are documented in refs 9, 10. The sulphur standard referred to is the Cañon Diablo meteorite troilite; sulphur isotope values have an analytical uncertainty of ~0.05‰. All isotopic analyses were made on two V. G. Micromass 602-C mass-spectrometers.

Brief sample descriptions and isotopic data are presented in Table 1. Sulphur fractionation between seawater, evaporite sulphate and baryte is small; therefore the low (10–12‰) $\delta^{34}\text{S}$ values characteristic of Upper Permian seawater¹¹ can indicate derivation of sulphate within both evaporites and baryte. Interbedded displacive evaporite nodules and replacive evaporite crystal clusters from core sample (Table 1, S1–2) both gave values typical of Upper Permian seawater. The replacive evaporites thus either grew during the Upper Permian, or were formed by later remobilization of evaporites of that age. Petrographic evidence¹² supports the former hypothesis.

In former displacive anhydrite nodules, calcite occurs as pseudomorphs of anhydrite cleavage flakes and, in places, contains corroded anhydrite inclusions². The nodules now have a narrow (<1 cm) margin of ferroan dolomite which grades outwards into host, non-ferroan dolomite. The ferroan dolomite is partially dedolomitized and contained in poikilitic calcite.

* Present address: Department of Geology, Louisiana State University, Baton Rouge, Louisiana 70803, USA.

Table 1 Results of isotopic analyses

Specimen no.	Description	$\delta^{34}\text{S}_{\text{CD}}$		
S1/152	Gypsum interbedded with dolomite	+11.40		
S2/164	Gypsum after replacive anhydrite	+11.46		
S3/463	Baryte in replaced anhydrite nodules	+17.88		
S4/463	Sphalerite in replaced anhydrite nodules	-21.92		
S5/463	Galena in replaced anhydrite nodules	-15.10		
S6/456	Marcasite in replaced anhydrite nodules	-29.44		
S7/456	Sphalerite in replaced anhydrite nodules	-15.98		
S8/673	Galena replacing dolomite nodules	-32.33		
S9/700	Baryte replacing fenestral dolomite	+11.21		
S10/619	Baryte in vug from dispersed baryte-sulphide assemblage	+17.34		
S11/304	Galena in mineralized breccia	-7.96		
S12/523	Galena in void fill in sucrose dolomite	-8.06		
S13/644	Galena with calcite in veinlet	-8.51		
		$\delta^{13}\text{C}_{\text{PDB}}$	$\delta^{18}\text{O}_{\text{PDB}}$	
C/01/603	Calcite spar from vug after dissolution of replacive anhydrite	-4.95	-7.27	
C/02/454	Calcite spar from vug after dissolution of displacive anhydrite	-2.64	-7.02	
C/03/463	Host dolomite surrounding replaced anhydrite nodules	+5.56	-2.31	
C/04/463	Ferroan dolomite in margin of replaced anhydrite nodules-separated from C/05	+5.20	-5.11	
		+5.36	-4.23	
C/05/463	Dedolomite (calcite) in margin of replaced anhydrite nodules	+3.53	-9.41	
C/06/463	Calcitized anhydrite from nodule centre	+3.24	-13.30	
C/07/604	'Rounded' calcite crystals from calcitized anhydrite nodules	+2.28	-13.57	

The sphalerite-galena-marcasite-baryte mineralization is intimately intergrown with calcite and apparently formed during calcitization of the anhydrite; both mineralization and calcitization were believed to be the result of bacterial reduction of the sulphate².

The isotopic compositions of the sulphides within these mineralized replaced nodules (Table 1, S4-7) suggest that growth took place in non-equilibrium conditions. In an equilibrium assemblage $\delta^{34}\text{S}$ would increase in the order galena-sphalerite-marcasite¹³. High negative $\delta^{34}\text{S}$ values are characteristic of sulphate reduction. Petrographic relations suggest that mineralization took place early in the sedimentary history, whilst still near surface with consequent low temperatures during mineralization. Abiotic sulphate reduction only readily takes place at temperatures over 200 °C (ref. 14). Desulphurization of hydrocarbons, possibly a low temperature reaction, produces little isotopic fractionation¹⁵ and the small amounts of sulphide produced by this process might have values between 0‰ and +5‰; thus bacterial reduction of the evaporite sulphate here appears the most probable mechanism. The large negative fractionation caused by this reduction may leave the remnant sulphate correspondingly enriched in ^{34}S ; the isotopic value of the baryte from the same replaced nodule (Table 1, S3), together with probable inclusions of anhydrite (ref. 2, Fig. 8), suggest such enrichment did take place.

Carbon and oxygen isotope values of calcite from the replaced nodules are distinct from those of the calcite spar found partially occluding vugs formed by near-surface dissolution of both replacive and displacive evaporites (Table 1, C/0 1-2). The isotope values of the calcite spar, with light $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ are what would be expected with calcite precipitation from Quaternary or Recent ground waters with the negative carbon values

Table 2 Calculated oxygen isotope temperatures

Specimen no.	Description	Calculated temperatures (°C)	
		$\delta^{18}\text{O}_{\text{H}_2\text{O}} = 0\text{‰}$	-5‰
C/03/463	Host dolomite	46	18
C/04/463	Ferroan dolomite	58	28
C/04a/463		65	33
C/05/463	Dedolomite (calcite)	69	38
C/06/463	Calcitized anhydrite	98	62
C/07/604	'Rounded' calcitized anhydrite	100	63

resulting from atmospheric and soil-derived carbon dioxide. The $\delta^{13}\text{C}$ values of the carbonates associated with the replaced anhydrite nodules (Table 1, C/0 3-7) are considerably higher: the heaviest values are typical of evaporitic environments. $\delta^{18}\text{O}$ values in carbonates are dependent on the origin of the precipitating pore fluids and the temperature of mineral formation. Original marine pore fluids would have had a value close to 0‰_{SMOW} or higher, dependent on the extent of evaporation. Both addition of isotopically light water from a meteoric source and increased temperature during crystallization would lead to a depletion of ^{18}O .

Evaporite sulphate reduction is more likely to take place in relatively sulphate-poor pore fluids, either diagenetically modified marine waters or waters with a meteoric addition. Carbon isotope fractionation has little dependence on temperature¹⁶ but isotopically light carbon can be introduced from oxidation of organic material *in situ* or from meteoric water. Diagenetic carbonates associated with bacterial sulphate reduction usually possess high negative values¹⁷. The slight decrease in $\delta^{13}\text{C}$ values here could result from buffering of bacterial carbon by carbonate-rich pore fluids.

The carbonate isotope data plot along a line (Fig. 1). This line may be a partial artefact from contamination during separation of the ferroan dolomite and dedolomite calcite (Table 1, C/0 4-5); however, the trend is still defined by the remaining samples (Table 1, C/03, 6-7) with both $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ becoming more negative with calcitization of anhydrite. The data represent mixtures of two carbonate sources, the heavier one evaporitic and the lighter one from bacterial sulphate reduction in water with a meteoric component (possibly at a higher temperature).

Maximum temperatures of crystallization of the carbonates were calculated using the equation of Craig¹⁸ and one calculated from the data of Fritz and Smith¹⁹. Table 2 gives the calculated temperatures for formation of the carbonates associated with the replaced anhydrite nodules for pore fluids with initial $\delta^{18}\text{O}_{\text{H}_2\text{O}}$ values of both 0 and -5‰_{SMOW}: the latter is a plausible minimum value for meteoric water in an evaporitic climate. The resultant low temperatures (Table 2) are further evidence for bacterial reduction of the sulphate.

The epigenetic baryte which replaces dolomite (Table 1, S9) has a $\delta^{34}\text{S}$ value very close to that of the evaporites. The sulphate in the baryte is thus thought to be derived directly from evaporite sulphate although, because of preservation of fenestral fabric within the dolomite^{1,4} it is unlikely that the dolomite was replaced by calcium sulphate before baryte replacement.

The sulphur isotope value from baryte in the epigenetic interspersed baryte-sulphide assemblage (Table 1, S10) is distinct from the evaporite sulphate value and represents diagenetically modified Upper Permian seawater sulphate. As freshwater sulphate is commonly much lighter than marine sulphate²⁰ there seems little likelihood of a dominant meteoric influence.

Apart from penecontemporaneous replacive galena near the base of the formation (Table 1, S8) other Upper Permian galenas have $\delta^{34}\text{S}$ averaging -8‰ (Table 1, S11-13); their origin is therefore distinct from the bacteriologically reduced sulphides, amongst which S8 can be included. These sulphur

isotope values may result from partial reduction of Upper Permian seawater sulphate within a closed system; however, the apparent consistency of values from widely spaced localities suggests a common, but unknown, origin.

The isotopic data in this study support most of the inferences drawn by Harwood² and Harwood and Smith¹. Both displacive and replacive evaporites have $\delta^{34}\text{S}$ sulphate values characteristic of Upper Permian seawater. Where modification of pore fluid composition caused replacement and mineralization of anhydrite nodules, light oxygen values of the carbonates suggest either an increase in temperature before crystallization, or the introduction of meteoric water; a combination of both these factors is more probable. Meteoric water could also have caused the associated slight depletion in $\delta^{13}\text{C}$. However, the variable and very light non-equilibrium compositions of the sulphides suggest bacterial reduction on an organic substrate and is a better explanation of the isotopically lighter carbon.

We thank NERC under whose auspices the isotopic analyses were made possible, and we thank Miss Pearline Raymond for typing this paper.

Received 8 September; accepted 13 December 1982.

1. Harwood, G. M. & Smith, F. W. *Proc. EZ82 Conf. Spec. Vol. geol. Soc. Lond.* (in the press).
2. Harwood, G. M. *Contr. Sedim.* **9**, 61–72 (1980).
3. Deans, T. *Miner. Mag.* **32**, 705–715 (1961).
4. Harwood, G. M. thesis, Open Univ. (1981).
5. McCrea, J. M. *J. Chem. Phys.* **18**, 849–857 (1950).
6. Craig, H. *Geochim. cosmochim. Acta* **12**, 133–149 (1957).
7. Deines, P. *Int. J. Mass Spectrom. Ion Phys.* **4**, 283–295 (1970).
8. Sharma, T. & Clayton, R. N. *Geochim. cosmochim. Acta* **29**, 1347–1353 (1965).
9. Robinson, B. W. & Kusakabe, M. *Analyt. Chem.* **47**, 1179–1181 (1975).
10. Coleman, M. L. & Moore, M. P. *Analyt. Chem.* **50**, 1594–1595 (1979).
11. Claypool, G. E., Holser, W. T., Kaplan, I. R., Sakai, H. & Zak, I. *Chem. Geol.* **28**, 199–260 (1980).
12. Harwood, G. M. *Proc. EZ82 Conf. Spec. Vol. geol. Soc. Lond.* (in the press).
13. Rye, R. O. & Ohmoto, H. *Econ. Geol.* **69**, 826–842 (1974).
14. Dhannoun, H. Y. & Fyfe, W. S. *Prog. Expl. Petrol., N.E.R.C. Publ. Ser. D.* **69**–71 (1972).
15. Orr, W. L. *Bull. Am. Ass. petrol. Geol.* **50**, 2295–2318 (1974).
16. Emrich, K., Erhart, D. H. & Vogel, J. C. *Earth planet. Sci. Lett.* **8**, 363–371 (1970).
17. Hudson, J. D. *J. geol. Soc. Lond.* **133**, 637–660 (1977).
18. Craig, H. in *Stable Isotopes in Oceanographic Studies and Palaeotemperatures* (ed. Tongiorgi, E.) 3–24 (Consiglio Nazionale delle Ricerche, Pisa, 1965).
19. Fritz, P. & Smith, D. G. W. *Geochim. cosmochim. Acta* **34**, 1161–1173 (1970).
20. Coleman, M. L. *J. geol. Soc. Lond.* **133**, 593–608 (1977).

Alkylbenzene pollution of Tokyo Bay sediments

Ryoshi Ishiwatari, Hideshige Takada & Sun-Ja Yun

Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, Fukazawa, Setagaya-ku, Tokyo 158, Japan

Eiji Matsumoto

Marine Geology Department, Geological Survey of Japan, Yatabe, Tsukuba-gun, Ibaraki-ken 305, Japan

Annual discharges of organic matter into Tokyo Bay have been estimated to be $\sim 1.5\text{--}6.6 \times 10^5$ tonnes (expressed as carbon) during 1973–76¹. We report here new findings of alkylbenzene pollution of sediments from the bay and adjacent areas. The pollution level of alkylbenzenes (0.5–1.4 μg per g dry sediment) is comparable with that of polycyclic aromatic hydrocarbons. The vertical distribution of alkylbenzenes in a Tokyo Bay sediment indicates that the pollution started around 1960—at the same time as alkylbenzenesulphonates (synthetic detergents) began to be used widely in the Tokyo area. The alkylbenzenes in the sediments are considered to have been carried with alkylbenzenesulphonates.

In the course of analysing polycyclic aromatic hydrocarbons (PAH) in sediment samples from one of the major polluted rivers (Tamagawa River) flowing into Tokyo Bay, we found

Table 1 Relative abundances of homologues of alkylbenzenes and alkylbenzenesulphonates (ABS) from the Tamagawa River sediments

Compound	Sampling location	Concentration (μg per g dry sediment)	Relative abundance (% of the total alkylbenzenes or ABS)				
			C ₁₀	C ₁₁	C ₁₂	C ₁₃	C ₁₄
Alkylbenzenes (this study)	Rokugo	6.4	4	25	35	30	6
ABS (ref. 4)	Rokugo	86.3	ND	35	48	16	ND
	Maruko	49.5	3	27	39	30	2

ND, not detected.

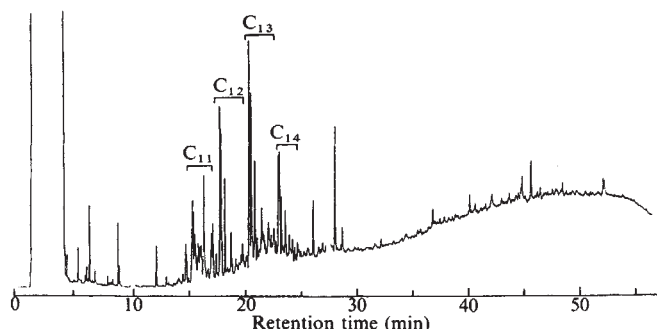


Fig. 1 A representative gas chromatogram of alkylbenzenes from a Tokyo Bay sediment. This was obtained using a Hewlett-Packard 5880A gas chromatograph equipped with a FID detector. A 0.3 mm i.d. $\times$ 25 m fused silica capillary column coated with SE-54 was used for the analysis in the following conditions: oven temperature 50 °C (maintained for 2 min) to 120 °C at 30 °C min⁻¹ and then to 190 °C at 3 °C min⁻¹; a helium flow rate 1.5 ml min⁻¹; splitless mode.

that some unknown compounds were present in an *n*-hexane fraction which was expected to contain hydrocarbons with more aliphatic than PAH. Gas chromatography indicated the presence of more than 20 compounds in the fraction (Fig. 1). The electron impact mass spectra taken inserted with a Shimadzu LKB 9000 gas chromatograph-mass spectrometer (GC-MS) for these compounds indicated intense ions at m/z 91, 105, 119 and/or 133 with less abundant ions at m/z 147, 161, 175. These mass spectra were clearly characteristic of alkylbenzenes. The molecular ions appeared at m/z 218, 232, 246, 260, 274 or 288 indicated the alkyl chain lengths of these alkylbenzenes to be 10, 11, 12, 13, 14, and 15, respectively. Furthermore, by comparison with standard mass spectra², we tentatively assigned these alkylbenzenes as: 3- and 4-phenyl decanes (3-C₁₀ and 4-C₁₀); 2-, 3-, 4-, 5- and 6-C₁₁; 2-, 3-, 4-, 5- and 6-C₁₂; 2-, 3-, 4-, 5-, 6- and 7-C₁₃; 5-, 6- and 7-C₁₄ and phenyl heptadecanes (position of substitution is unknown). The molecular range of these alkylbenzenes and the relative abundance of their alkyl homologues are similar to those of alkylbenzenesulphonates (ABS) reported for a sediment sample from the same river (Table 1)³. This fact suggests that the alkylbenzenes in the river sediment are of the same origin as ABS used as a detergent.

To find evidence in favour of the above suggestion, we looked for a historical record of alkylbenzenes in a Tokyo Bay sediment. A core sediment sample was obtained from Tokyo Bay (Location no. G80-8: 35°32.5' N; 139°55.0' E). The sedimentation rate at the site was measured by a ²¹⁰Pb method⁴. Figure 1 shows a typical gas chromatogram of alkylbenzenes and C₁₁–C₁₄ alkylbenzenes are clearly seen. Table 2 gives the absolute concentrations and relative abundances of the alkyl homologues of alkylbenzenes from this core sample. As the concentrations of alkylbenzenes are extremely low for the sections below 30 cm in depth and thus the errors in the concentrations are considered to be large, we do not give the relative abundance in Table 2. Two marked features of these data can

Table 2 Concentrations and relative abundance of alkylbenzenes in a sediment core from Tokyo Bay (35°32.5' N; 139°55.0' E)

Depth (cm)	Approximate year of deposition*	Concentration of alkylbenzenes (ng per g dry sediment)					Total
		C ₁₀	C ₁₁	C ₁₂	C ₁₃	C ₁₄	
0–5		0	67	170	276	79	592
		(0)†	(11)	(29)	(47)	(13)	
5–10		8	201	369	533	187	1,298
		(1)	(16)	(29)	(41)	(14)	
10–15	1970	2	117	205	318	147	789
		(0)	(15)	(26)	(40)	(19)	
15–20		3	78	174	265	114	633
		(0)	(12)	(28)	(42)	(18)	
20–25		11	172	380	585	235	1,383
		(1)	(12)	(28)	(42)	(17)	
25–30	1960	7	101	144	167	101	520
		(1)	(19)	(28)	(32)	(19)	
30–35		0	4	27	24	0	55
35–40	1950	0	12	7	0	0	19
40–45		0	0	0	0	0	0
45–50	1940	0	0	0	0	0	0

A typical procedure for separating alkylbenzenes from a sediment is as follows: a powdered freeze-dried sediment sample was Soxhlet-extracted in a glass-fibre thimble with benzene-methanol (6:4) for 8 h. The extract was concentrated to nearly dryness in a rotary evaporator at 30 °C. Elemental sulphur was then removed from the extract by percolating through an activated copper column (0.8 cm i.d. × 5 cm). Benzene-methanol (6:4) eluate was evaporated to dryness, taken up 1 ml of benzene and then subjected to percolated through a Florisil column (1 cm i.d. × 5 cm). A benzene eluate was evaporated to dryness, taken up 0.3 ml of *n*-hexane and subjected to silica gel column chromatography (Mallinckrodt 100 mesh; 0.5 cm i.d. × 18 cm). The column chromatography was performed under a slight pressure of nitrogen, giving a flow rate of ~0.2 ml min⁻¹. The same solvent (*n*-hexane) was used to elute two fractions (the first of 6 ml, the second of 10 ml). The second *n*-hexane fraction contained alkylbenzenes. This fraction was subsequently analysed by gas chromatography and gas chromatography-mass spectrometry.

* Estimated by a Pb-210 method⁵.

† Relative abundance expressed as % of the total alkylbenzenes.

be noted: (1) The absolute concentrations (0.5–1.4 µg per g dry sediment) of alkylbenzenes in the sections from surface to 30 cm in depth (younger than 1960) greatly exceeded those in the deeper sections. (2) The relative abundances of alkylbenzenes are nearly constant in the 0–30 cm sections: C₁₀:C₁₁:C₁₂:C₁₃:C₁₄ = 1 ± 1:4 ± 3:28 ± 1:41 ± 5:17 ± 3. This fact suggests that the nature of source materials has been constant since ~1960.

The abrupt increase of the absolute concentrations of alkylbenzenes in the core corresponds well to the rapid increase in the production of synthetic detergents in Japan, which have started up since 1958. The vertical distribution of alkylbenzenes does not resemble that of PAH for the same core sample: PAH are present in a fair amount (0.7–1.2 µg per g dry sediment) even in the 30–50 cm sections (1.4–2.0 µg per g for the upper sections). This fact indicates that the alkylbenzenes are not directly related to the pollution of oil or fossil fuel combustion materials. We conclude that the alkylbenzene pollution is linked with the use of ABS detergents around the Tokyo area. This conclusion is in good agreement with the work of Ambe⁵ on Tokyo Bay sediments. He found that ABS pollution occurs only in the upper sections (30 or 70 cm) of the bay sediments. Pollution level is 1.5–2 µg per g dry sediment.

One possible genetic relation between alkylbenzenes and ABS is the microbial transformation of ABS into alkylbenzenes in sediments. However, no studies so far conducted support this possibility. In general, microbial attack on alkyl chains of ABS takes place much more easily than microbial desulphonation⁶. Another possibility is that alkylbenzenes have been

carried with ABS into sediments. Since ABS are synthesized from benzene and olefins through alkylbenzenes as intermediate substances⁷, it is likely that a significant amount of alkylbenzenes are carried into ABS (final products). This was evidenced by the finding of C₁₀–C₁₄ alkylbenzenes (1.3% of ABS) in a ABS specimen (supplied by Lion Fat & Oil Co., Japan).

It is interesting from the environmental chemical point of view, that the ratio of concentration of ABS (data from ref. 5) to that of alkylbenzenes (this work) in Tokyo Bay is ~1. This ratio is much lower than that for the synthetic ABS (~100 as described above). This fact suggests that ABS molecules are broken down more easily than the closely related long chain alkylbenzenes, perhaps because the detergent molecules are water soluble and hence can be assimilated into biological system whilst the very insoluble alkylbenzenes cannot. Consequently, alkylbenzenes might be a better tracer of detergent ABS pollution than ABS themselves and the ratio of ABS to alkylbenzenes might indicate the degree of degradation of ABS in the aquatic environment.

This work was partly supported by the Ministry of Education, Science and Culture, Japan (grant no. 57035050).

Received 13 September; accepted 6 December 1982.

1. Ogura, N. *Bull. Coastal Oceanogr.* **16**, 92–100 (1979).
2. Stenhagen, E., Abrahamsson, S. & McLafferty, F. W. (eds) *Registry of Mass Spectral Data* Vols 2, 3 (Wiley, New York, 1974).
3. Hon-nami, H. & Hanya, T. *Jap. J. Limnol.* **41**, 1–4 (1980).
4. Matsumoto, E. *J. geol. Soc. Jap.* **81**, 301–306 (1975).
5. Ambe, Y. *Nature* **239**, 24–25 (1972).
6. Gledhill, W. E. *Adv. appl. Microbiol.* **17**, 265–293 (1974).
7. *Encyclopedia Genre Japonica* Vol. 15, 390 (Shogakukan, Tokyo, 1974).

Avoidance of neonatal cortical lesions by developing somatosensory barrels

Muneyuki Ito & Misako L. Seo

Institute for Developmental Research, Aichi Colony for the Handicapped, Kamiya-Cho, Kasugai 480-03, Japan

The barrel-shaped cellular aggregates in layer IV of the rodent somatosensory cortex have been used as landmarks to study the normal and abnormal development of cortical connection. Although much is known about the pattern of barrel field organization following early damage to the sensory periphery, it is unknown whether a neonatal lesion causes a defect in the barrel field or the lesion becomes circumvented by developing barrels. This issue is relevant to the problem of locus specificity versus plasticity of sensory cortex in very young animals. We report here that a focal lesion made in the somatosensory cortex soon after birth did not produce a defect in the barrel field in the adult. Rather, all the barrels seemed to avoid the lesion. Moreover, there was a strong tendency for the barrels to develop in a tight relationship so that the lesion was usually found outside the entire barrel field, as though it had been pushed aside.

The barrels in the rodent somatosensory cortex have been shown to represent whiskers on the contralateral side, and this 'one whisker-one barrel' correspondence is one of the rare cases in which the sensory periphery can be visualized on the cortex using common histological techniques¹, and has been used as a landmark to investigate cerebral neuronal plasticity during maturation. If a row (or rows) of whiskers is removed soon after birth, the cortical area normally connected to these whiskers becomes associated with the remaining whiskers^{2–4}, suggesting that the newborn cortical areas are not yet committed

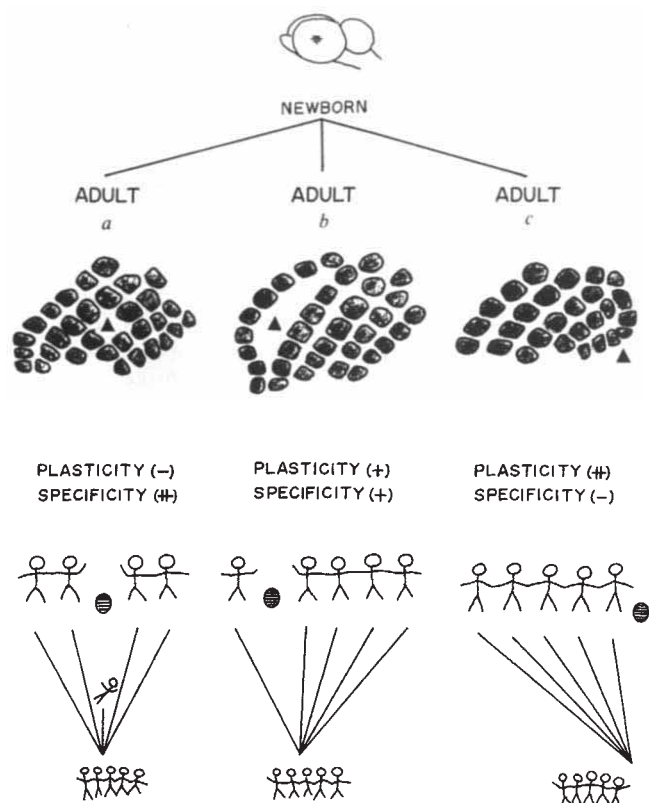


Fig. 1 At the top are shown three possible patterns of mature barrel fields, with emphasis on their relationship to a localized cortical lesion (triangle) made neonatally. *a*, At birth, each cortical point is already allocated to a whisker on the contralateral snout. The damaged locus remains as a defect in the barrel field. *b*, The lesion is circumvented by developing barrels. The thalamocortical fibres originally destined for the lesioned part of the cortex grow into the neighbouring intact area. The locus specificity is not absolute but relative. *c*, Even if the lesion is made within the prospective barrel area of neonate cortex, it is circumvented by the whole set of barrel-bound thalamocortical fibres, so that the lesion is found outside the entire barrel field, as though pushed aside. Bottom, the appearance of the barrel field is compared with schoolchildren trained to fan out to form a line; one morning they find an unexpected hole in the ground. With developing barrels, in practice, possibility *a* never occurred (as far as cortical lesioning is concerned). Some cases were consistent with possibility *b*, but most suggested possibility *c*.

as to which whisker they represent. To test this hypothesis more directly, we placed a localized lesion in the newborn somatosensory cortex, then examined the morphology of the barrel field in the adult. Figure 1 summarizes our working hypothesis: if every point on the cortical surface is already allocated a whisker at birth, a neonatal lesion should later result in a defect in the barrel field (Fig. 1*a*). If, on the other hand, the newborn cortex has some plasticity, all the barrels should appear on the cortex (Fig. 1*b, c*).

Wistar rats were used in the present study. Within 24 h after birth, the rats were anaesthetized with Nembutal (25 µg per g body weight) and placed on a stereotaxic headholder. A hypodermic needle, insulated except for 0.5 mm at the tip, was passed into the somatosensory cortex through a burr hole in the skull. Preliminary experiments showed that advancing the electrode 0.8 mm below the pial surface 0.5 mm anterior to bregma and 4.0 mm lateral to the midline brought the tip to the depth of the 'future' lamina IV of the whisker-barrel area (some dimpling of the cortical surface could not be avoided). A positive d.c. current of 200 µA was passed for 20 s. The

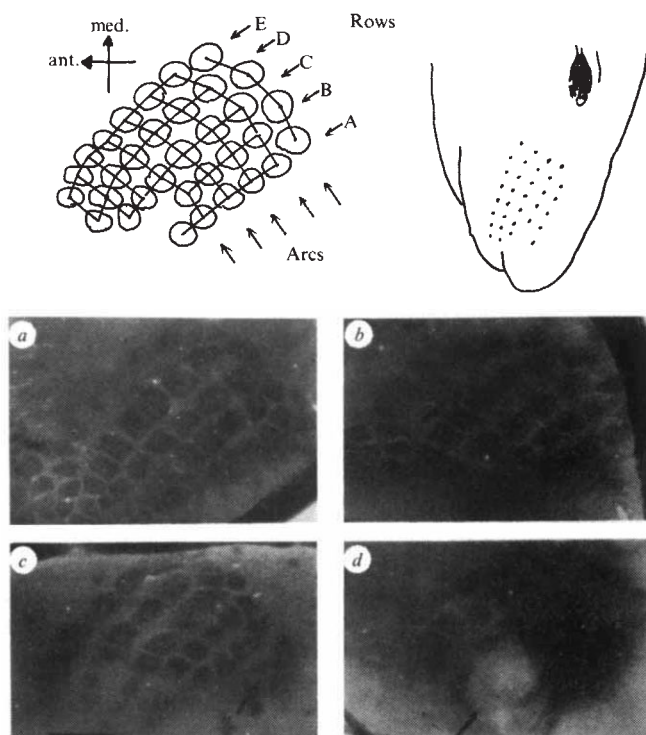


Fig. 2 The diagram at the top explains the correspondence between the whiskers in the periphery and the barrels in the cortex. The terms used to specify aspects of the barrel field are also indicated. At the bottom are shown photomicrographs of barrel fields. *a*, Barrel field in a normal control cortex; *b*, a neonatal cortical lesion was circumvented by rows D and E as indicated by the remnant of the lesion (arrow) and by widening of the septum between these rows. In general, lesions in SDH sections were not as clearly recognizable as in Nissl-stained sections, as the scar could not be identified as such in the former. However, starting from a superficially-recognizable lesion spot we could easily follow a small star-like hole in successive sections during tangential cutting. *c*, Example of commonly observed barrel fields after cortical lesioning at birth. The entire barrel field appears to keep in close contact, thus the lesion seems to have been pushed aside. Camera lucida drawings from examples *b* and *c* were used as models *b* and *c* of Fig. 1. *d*, When thalamocortical fibres were damaged subcortically within 24 h of birth, the corresponding barrels never developed. This indicates that each neonate thalamic fibre is already committed with respect to which whisker it represents. Scale bar in *c* represents 1 mm and applies to *a*, *b* and *d* also.

lesion thus made was circular with a diameter of ~0.4 mm as determined in animals killed immediately after lesioning, and it was recognized as a reddened spot macroscopically and as a scar of ~0.3 mm diameter in Nissl-stained coronal sections in adult animals examined. At 3 months of age the animals were given an overdose of Nembutal. The brain was removed from the skull and trimmed so that it could be sectioned in a plane tangential to the whisker field of the somatosensory cortex. Frozen sections (100 µm) were cut and collected in a saline-filled Petri dish. The saline was then replaced by 20 ml of incubation solution for succinic dehydrogenase (SDH) histochemistry^{5,6} and the specimen incubated at 37 °C for 20 min.

The assessment of the effect of the neonatal cortical lesion on barrel field formation was based on SDH data obtained from 23 cortically lesioned and 9 unoperated normal control rats from 11 litters. The barrel fields of normal rats were similar, in several respects, to those reported previously^{1,2,5,7-11}. The area occupied by the main barrels (corresponding to whisker rows A, B, C, D and E) was ~4.0 mm × 2.5 mm in tangential section. A given barrel was visible in 2-3 successive 100-µm

Table 1 Barrel field areas in normal and neonatally lesioned cortex of rats

	Normal cortex	Lesioned cortex	
		Type <i>b</i>	Type <i>c</i>
<i>n</i>	9	7	15
Area ($\pm$ s.d.)	8.45 $\pm$ 0.89	7.70 $\pm$ 0.58	6.77 $\pm$ 1.18

Areas are expressed as mm². Areas for type *b* include lesioned and widened septum. The difference between normal and type *c* was significant at $P < 0.01$ (*t*-test).

thick sections. The barrels were oval or quasi-rectangular in these sections (Fig. 2a).

Figure 2b shows a case in which a somatosensory cortical lesion was made immediately after birth and the brain examined 3 months later. Rows D and E appear to have developed separately. The growing thalamocortical fibres apparently skirted the lesion which was made between the prospective rows D and E (the original lesion was usually located as a star-like cavity in SDH sections). Thus, the neonatal lesion did not cause a defect in the barrel field, nor was there any indication that the barrels apposing the lesion were any smaller than those of corresponding controls.

This effect was seen in 8 of the 23 animals which received a neonatal cortical lesion. This result is consistent with model *b* of Fig. 1. Note also that a lesion, if it falls within the barrel field at all, always falls between two rows but not between two arcs⁹. Barrels in a row develop hand in hand. In these eight animals the avoidance of the lesion occurred between rows C and D (four cases) or between rows D and E (four cases). The lesion was always aimed at the centre of the vibrissa field, thus it can hardly be due to chance that no avoidance was evident between rows A, B and C. Obviously, rows A, B and C tend to develop in close relationship.

In the remaining 15 cortically lesioned rats the barrel appeared normal: there was no distortion of rows of barrels or any defect in the barrel field (Figs 1c, 2c). The entire barrel field avoided the neonatal lesion in these cases. One might argue that the lesion did not 'hit' the future barrel field. However, a neonatal lesion made 2.0 mm below the pial surface could easily hit the subcortical structure underlying the barrel field (see below). Therefore, we believe that even if a lesion is made in the future barrel field, if slightly caudal to the field centre, the barrels appear rostral to the lesion which itself moves more caudally during maturation and vice versa. This was the reason why, in the preliminary experiments, at first it seemed almost impossible to determine the coordinate of a neonatal cortical area which would be occupied by the barrels in the adult. Although the barrel fields in these cases were similar to those in the unoperated controls in overall organization, the total area of the field for five main rows of large mystacial vibrissae was significantly smaller than that from the normal controls (6.77 ± 1.18 mm² (mean $\pm$ s.d., $n = 15$) compared with 8.45 ± 0.89 mm² ($n = 9$); $P < 0.01$, *t*-test, see Table 1).

Circumvention of a lesion by the developing barrels may have resulted from one of two mechanisms: either the thalamocortical fibres grow inwards, skirting the lesion to terminate on intact neighbouring cortical cells, or the thalamocortical fibres grow straight upwards but reorganization occurs at the thalamic level, as thalamic afferent terminals within the cortex (if already there) may also be affected by the lesion and hence retrograde degeneration may reach thalamic relay cells at the origin. If the latter were the case, a subcortical lesion made at birth should also yield a barrel field of type *b* or *c* of Fig. 1; this was tested in seven animals derived from two additional litters. Using the same coagulating current and coordinates on the skull as for cortical lesioning, neonatal lesions were made via an electrode placed subcortically (2.0 mm below the pial surface). All seven animals were examined by the SDH

histochemical procedure after maturation and showed a defect in the barrel field (Fig. 2d). Thus, the thalamus is probably not the site of plastic changes responsible for modification of barrel fields in the cortically operated group. More importantly, the difference in the barrel field pattern between the cortically and subcortically operated groups indicates that the thalamocortical afferent terminals are not present in the cortex at birth, consistent with the view of Lund and Mustari¹².

A given point on the newborn rat vibrissa cortex is not strictly allotted to any particular whisker on the contralateral snout. However, a certain degree of topographical specificity is also apparent: the fact that the remnant of the neonatal lesion was often found between two barrel rows implies that the site of lesioning had originally been allotted to a whisker belonging to either of the two rows. Thus, two seemingly contrasting mechanisms, locus specificity and plasticity, are operating simultaneously. Attraction by a cytochemical substance has been proposed as a possible means of determining the direction and the site of termination of ingrowing fibres^{13,14}. The chemical nature or concentration of the chemical substance may be varied continuously over the cortical area. In view of the present findings the chemo-specificity is not absolute: if a cortical locus is damaged, developing fibres originally destined for that locus now 'seek' a neighbouring area associated with a slightly different substance or different concentration.

At the start of the present study we performed electrophysiological mapping of the somatosensory cortex. Using the recording method described previously⁷, 12 normal and 13 cortically lesioned animals were examined by repeated microelectrode penetrations at 0.5-mm intervals over the somatosensory cortex. However, as no difference was found between the two groups, we decided to use the SDH method. Probably the interval of 0.5 mm adopted in the physiological mapping provided insufficient resolution for detecting subtle differences (for example, widening of the septum or reduction in the size of the whisker receiving area) that could only be revealed by anatomical methods.

It was remarkable that during development, the main rows of barrels maintained a close relationship. This is in keeping with the concept of 'systems-matching' proposed by Gaze and Keating¹⁵ as a mechanism operating during formation of neuronal connections. Apparently, the whole set of whisker-related thalamic afferents squeeze into a limited space of somatosensory cortex, displacing the lesion. This view was substantiated by the finding of a significant difference in the total area of the barrel field between the cortically lesioned type *c* cases and the control groups. As barrels have been shown to appear 3–5 days postnatally⁹ and removal of whisker follicles before this critical period can produce changes in the barrel field configuration, it is of prime importance to determine how soon after birth a cortical lesion has to be made for the developing barrels to avoid the lesion. It would then be possible to determine whether peripheral and central lesioning can be equated as regards plasticity of the developing central nervous system.

This work was supported in part by grant 56570057 from the Ministry of Education in Tokyo.

Received 31 August; accepted 6 December 1982.

1. Woolsey, T. A. & Van der Loos, H. *Brain Res.* **17**, 205–242 (1970).
2. Van der Loos, H. & Woolsey, T. A. *Science* **179**, 395–398 (1973).
3. Waite, P. M. E. & Taylor, P. K. *Nature* **274**, 600–602 (1978).
4. Weller, W. L. & Johnson, J. I. *Brain Res.* **83**, 504–508 (1975).
5. Killackey, H. P. & Belford, G. R. *J. comp. Neurol.* **183**, 285–304 (1979).
6. Nachlas, M., Tsau, K., DeSouza, G., Cheng, C. & Seligman, A. *J. Histochem. Cytochem.* **5**, 420–436 (1957).
7. Ito, M., Kawabata, M. & Shoji, R. *J. Neurophysiol.* **42**, 1711–1726 (1979).
8. Ito, M. *J. Neurophysiol.* **46**, 705–715 (1981).
9. Rice, F. L. & Van der Loos, H. *J. comp. Neurol.* **171**, 545–560 (1976).
10. Welker, C. *Brain Res.* **26**, 259–275 (1971).
11. Woolsey, T. A., Welker, C. & Schwartz, R. H. *J. comp. Neurol.* **164**, 79–94 (1975).
12. Lund, R. D. & Mustari, M. J. *J. comp. Neurol.* **173**, 289–306 (1977).
13. Attardi, D. G. & Sperry, R. W. *Expl Neurol.* **7**, 46–64 (1963).
14. Jacobson, M. & Gaze, R. M. *Expl Neurol.* **13**, 418–430 (1965).
15. Gaze, R. M. & Keating, M. J. *Nature* **237**, 375–378 (1972).

Facilitated induction of hippocampal long-lasting potentiation during blockade of inhibition

H. Wigström & B. Gustafsson

Department of Physiology, University of Göteborg, Göteborg, Sweden

A long-lasting potentiation of synaptic transmission can be induced in the hippocampus by high-frequency stimulation of the afferent fibres^{1,2}. This process has been regarded as a possible physiological substrate for long-term memory. From theoretical considerations on memory function it has been proposed that synaptic strengthening occurs as a result of simultaneous pre- and postsynaptic activation³, a principle which has been used in several theoretical models (for references see refs 4, 5). It thus seems important to study factors which control hippocampal long-lasting potentiation and, in particular, whether these factors act pre- or postsynaptically. A presynaptic control of long-lasting potentiation was established in experiments where two separate inputs to the same population of CA1 pyramidal cells were used^{6,7}. Studies on the granule cell population of area dentata however, revealed, a heterosynaptic modulation of the generation of potentiation, suggesting a possible postsynaptic control also^{8,9}. We have now studied long-lasting potentiation in the hippocampal slice preparation (CA1 region) in conditions where postsynaptic inhibition was reduced by application of γ -aminobutyric acid (GABA) receptor blockers. We found that in addition to their direct effect on inhibition, GABA blockers dramatically facilitated the induction of long-lasting potentiation. The results indicate involvement of postsynaptic mechanisms in the generation of long-lasting potentiation.

Experiments were performed on transverse hippocampal slices from guinea pigs ($n = 8$) using standard techniques¹⁰. CA1 pyramidal cells were synaptically activated by electrical stimulation of fibres in the apical dendritic layer (Fig. 1A, B), and the field potentials evoked in this layer and in the cell body layer were recorded. Higher than normal concentrations of calcium and magnesium ions (4 mM) were used because it reduced the risk of seizure activity when inhibition was blocked. No noticeable effect on long-lasting potentiation was seen using this solution, which is therefore referred to as normal. The GABA blockers picrotoxin, bicuculline and penicillin were administered via the perfusion solution or by means of droplets applied to the surface of the slice.

To generate long-lasting potentiation in the slice preparation, tetanization with 100 or more impulses is normally used^{2,6,7}. In the present series of experiments considerably shorter tetani were used, consisting of only 5–10 impulses. Possible effects of these minimal tetani were studied in the absence and presence of GABA blockers, and typical results are shown in Fig. 1C. The left column displays recordings from the dendritic layer, showing the initial presynaptic volley and the following field excitatory postsynaptic potential (e.p.s.p.), the slope of which is used as a measure of synaptic efficacy. The right column gives recordings from the cell body layer, showing the population spike (absent in some of the recordings) which is an indicator of the amount of pyramidal cell firing.

Figure 1Ca shows recordings taken after an initial 15-min control period demonstrating stable responses. The stimulus was in this case below strength for evoking a population spike. A brief tetanus (10 pulses at 50 Hz) was then delivered. This produced only a slight short-lasting (<30 s) increase in the potentials. Recordings taken 5 min after tetanization are shown in Fig. 1Cb (solid curve) together with the replotted control response (dotted), demonstrating complete absence of long-lasting potentiation. After this test a small drop of picrotoxin

solution was applied to the slice. This resulted in the appearance of a population spike in the recording from the cell body layer (Fig. 1Cc, right) probably due to blockade of feedforward inhibition¹¹. A reversed version of the spike is seen in the dendritic recording superimposed on the field e.p.s.p. (Fig. 1Cc, left). The rising phase of the field e.p.s.p. did not change, however, except for a small increase near the peak. The absence of changes in the initial portion of the field e.p.s.p. indicates that picrotoxin application did not influence excitatory synaptic transmission.

The effect of another brief tetanus, now delivered to the picrotoxin-treated slice, is seen in Fig. 1Cd, comparing recordings taken 5 min after tetanization (solid) with pre-tetanic ones (dotted). The slope of the rising phase of the field e.p.s.p. is increased (35%), indicating enhanced excitatory synaptic transmission. The population spike also increased and even a second one occurred (not visible in the figure). Following a slight decrease, the field e.p.s.p. attained a steady level of 30% increase which was present throughout the 30-min post-tetanic recording period.

In 18 of 20 slices, short tetani (10 impulses) which did not produce long-lasting potentiation in normal conditions did so after application of picrotoxin. The average field e.p.s.p. increase (slope) was 27% at 15 min after tetanus. Similar results were obtained in slices treated with bicuculline (7/7) and penicillin (6/6). Furthermore, in addition to the facilitation of the induction of long-lasting potentiation, there seemed to be

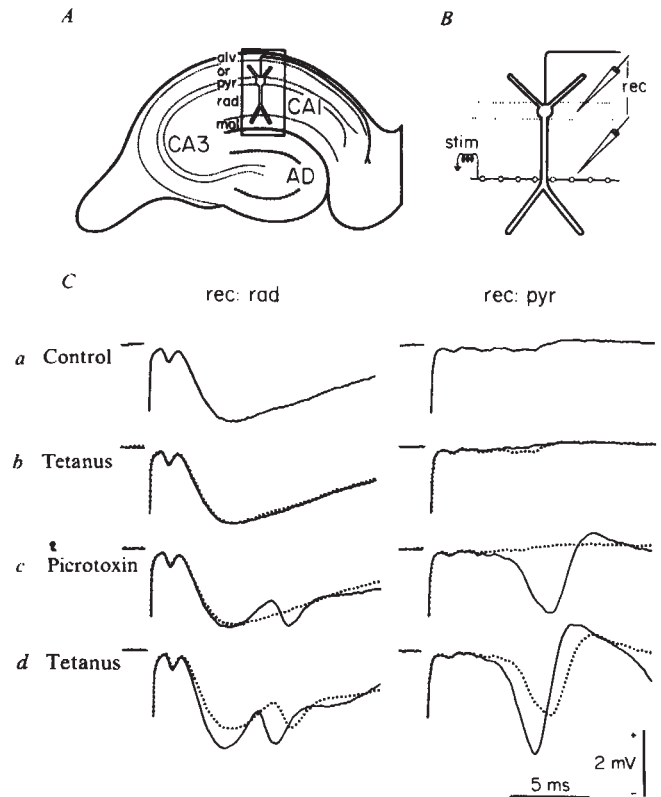


Fig. 1 Experimental arrangement (A, B) and the effect of brief tetanization before and after picrotoxin application (C). A, schematic drawing of the hippocampal slice with the studied CA1 region boxed in. Abbreviations: alv, alveus; or, stratum oriens; pyr, stratum pyramidale; rad, stratum radiatum; mol, stratum moleculare. B, diagram showing the arrangement of stimulating (stim) and recording (rec) electrodes. C, extracellular responses from apical dendritic layer, stratum radiatum (left column) and cell body layer, stratum pyramidale (right column). Each trace represents an average of 10 successive recordings. a, Control responses. b, 5 min after a 10-impulse, 50-Hz tetanus (solid); a repeated (dotted). c, After application of a 10 mM droplet of picrotoxin near the recording area (solid); b repeated (dotted). d, 5 min after a 10-impulse, 50-Hz tetanus in the picrotoxin-treated slice (solid); c repeated (dotted).

some increase in the saturation level obtained after repeated tetanization.

The amount of pyramidal cell firing evoked by the stimulus increased during blockade of inhibition (Fig. 1Cc). This effect was very pronounced during tetanization, although the frequency potentiation of the field e.p.s.p. was unaffected. This might suggest that increased postsynaptic firing is involved in the facilitated induction of long-lasting potentiation. Although it is possible that such vigorous firing would produce a non-specific effect different from long-lasting potentiation seen in normal conditions, experiments with two inputs to the same dendritic tree (see ref. 6) showed, that long-lasting potentiation in the picrotoxin-treated slice is specific for the tetanized pathway, in accordance with long-lasting potentiation produced in normal conditions.

The three GABA blockers used are known to affect both presynaptic¹²⁻¹⁴ and postsynaptic^{15,16} inhibition. We believe that presynaptic inhibition has not previously been demonstrated in the hippocampus. Lack of presynaptic inhibition is also indicated by the observation that the GABA blockers did not influence excitatory synaptic transmission, either during single responses (Fig. 1Cc) or during tetanic activation (unchanged frequency potentiation). In contrast, the ability of GABA blockers to reduce postsynaptic inhibition in the hippocampus is well known¹⁷⁻¹⁹. It thus seems likely that the facilitated induction of long-lasting potentiation is in some way—directly or indirectly—related to this decrease in postsynaptic inhibition. As mentioned above, firing in the postsynaptic cell might act as a mediator. Earlier attempts to control firing by varying the postsynaptic membrane potential of single pyramidal cells have, however, failed to demonstrate an effect on long-lasting potentiation²⁰. Moreover, the fact that potentiation is specific for the tetanized input would at least exclude somatic firing as a mechanism, leaving dendritic firing as a more conceivable alternative. The ability of dendrites to generate action potentials has previously been reported for both CA1 pyramidal cells^{21,22} and dentate granule cells²³. Presumed dendritic population spikes were also observed in the present experiments during tetani that were effective in evoking potentiation. In view of recent findings of dendritic feedforward inhibition in hippocampal pyramidal cells¹¹ it seems possible that active generation of dendritic depolarization could be controlled by inhibition. Note in this connection that picrotoxin blockade of inhibition in the cerebellum enables probable dendritic calcium-mediated depolarizations to be produced in response to parallel fibre activation²⁴. Thus, it is possible that the condition for synaptic strengthening during long-lasting potentiation is a combination of presynaptic transmitter release and processes associated with a large postsynaptic dendritic depolarization.

This work was supported by the Swedish MRC (Projects 05954 and 05180) and Magnus Bergvalls Stiftelse.

Received 7 September; accepted 16 December 1982.

- Bliss, T. V. P. & Lomo, T. *J. Physiol., Lond.* **232**, 331-356 (1973).
- Schwartzkroin, P. A. & Wester, K. *Brain Res.* **89**, 107-119 (1975).
- Hebb, D. O. *The Organization of Behavior* (Wiley, New York, 1949).
- Longuet-Higgins, H. C., Willshaw, D. J. & Buneman, O. P. *Q. Rev. Biophys.* **3**, 223-244 (1970).
- Amari, S.-I. *Biol. Cybernet.* **26**, 175-185 (1977).
- Dunwiddie, T. & Lynch, G. *J. Physiol., Lond.* **276**, 353-367 (1978).
- Andersen, P., Sundberg, S. H., Sveen, O., Swann, J. W. & Wigström, H. *J. Physiol., Lond.* **302**, 463-482 (1980).
- Levy, W. B. & Steward, O. *Brain Res.* **175**, 233-245 (1979).
- Douglas, R. M., Goddard, G. V. & Rives, M. *Brain Res.* **240**, 259-272 (1982).
- Skrede, K. K. & Westgaard, R. H. *Brain Res.* **35**, 589-593 (1971).
- Alger, B. E. & Nicoll, R. A. *J. Physiol., Lond.* **328**, 105-123 (1982).
- Eccles, J. C., Schmidt, R. & Willis, W. D. *J. Physiol., Lond.* **168**, 500-530 (1963).
- Curtis, D. R., Duggan, A. W., Felix, D. & Johnston, G. A. R. *Brain Res.* **32**, 69-96 (1971).
- Davidoff, R. A. *Brain Res.* **36**, 218-222 (1972).
- Bisti, S., Isolf, G. & Strata, P. *Brain Res.* **28**, 591-593 (1971).
- Lothman, E. W. & Somjen, G. G. *Electroenceph. clin. Neurophysiol.* **41**, 337-347 (1976).
- Curtis, D. R., Felix, D. & McLennan, H. *Br. J. Pharmacol.* **40**, 881-883 (1970).
- Dingledine, R. & Gjerstad, L. *Brain Res.* **168**, 205-209 (1979).
- Alger, B. E. & Nicoll, R. A. *Brain Res.* **200**, 195-200 (1980).
- Wigström, H., McNaughton, B. L. & Barnes, C. A. *Brain Res.* **233**, 195-199 (1982).
- Schwartzkroin, P. A. *Brain Res.* **128**, 53-68 (1977).
- Leung, L. S. *Brain Res.* **158**, 219-222 (1978).
- Jefferys, J. G. R. *J. Physiol., Lond.* **249**, 16-18P (1975).
- Ekerot, C.-F. thesis, Univ. Lund (1982).

Oxygen uptake occurs faster than sodium pumping in bee retina after a light flash

M. Tsacopoulos, R. K. Orkand*, J. A. Coles, S. Levy† & S. Poitry

Experimental Ophthalmology Laboratory, 22, rue Alcide-Jentzer, 1211 Geneva 4, Switzerland and Department of Physiology, Centre Médical Universitaire, 1211 Geneva 4, Switzerland

When neurones are active there is an entry of Na^+ , which must subsequently be pumped out, and an increase in their oxygen consumption rate (Q_{O_2}). The Na^+ pump derives its energy from ATP, splitting it into ADP and P_i , and it has reasonably been proposed that the changes in concentrations of ATP, ADP and P_i lead to a stimulation of the O_2 consumption by the mitochondria and hence to a restoration of the stock of ATP¹⁻³. Here we present evidence suggesting that Q_{O_2} must be controlled differently in the retinal photoreceptor cells of the honeybee drone. Stimulation of drone photoreceptors with a flash of light causes an entry of Na^+ (ref. 4) and a transient increase in Q_{O_2} that indicates respiration of the right order of magnitude to provide ATP to pump the Na^+ out⁵. We report intracellular recordings of changes in intracellular sodium (Na_i^+) and potassium (K_i^+) in response to single light flashes and have compared the time course of extra oxygen consumption (ΔQ_{O_2}) with these ion changes and other indices of Na^+ pumping. We found that the time course of pumping seems to lag behind the time course of ΔQ_{O_2} . It follows that the mitochondrial respiration must be stimulated by some signal which is generated earlier than the rise in ADP produced by the Na^+ pump.

Experiments were done on slices of the compound eye of the drone, *Apis mellifera* ♂, superfused on one or both faces^{5,6}. Light flashes were delivered perpendicular to the cut surface and stimulated all the photoreceptor cells simultaneously. Drone photoreceptors respond to a flash of light with a depolarizing receptor potential^{7,8}, as shown in Fig. 1a. This phototransduction is an amplification process that begins with the absorption of a relatively small number of photons by the visual pigment, rhodopsin, and leads to an influx of Na^+ ions through the surface membrane and an efflux of K^+ ions. We expect that most of the energy involved in the response is associated with this final, highly amplified, step. Tsacopoulos and Poitry⁵ made indirect tests of this hypothesis. Removal of Na from the superfusate reversibly reduced ΔQ_{O_2} by 63%, and prolonged application of a high concentration of strophanthidin (0.5 mM) reduced ΔQ_{O_2} and slowed down its time course. In the present study, we made more direct measurements that allowed us, first, to estimate how much of ΔQ_{O_2} is necessary to provide ATP for the Na^+ pump, and second, to compare the time courses of ΔQ_{O_2} and Na^+ pumping.

Double-barrelled, intracellular, ion-selective microelectrodes⁹⁻¹¹ were used to record changes in Na_i^+ and K_i^+ resulting from single, brief (20-40 ms) light flashes. Figure 1b and c show the transient increase in Na_i^+ and the corresponding decrease in K_i^+ . Such light flashes also cause a marked transient decrease in P_{O_2} in the retina that can be measured with an oxygen microelectrode¹². The minimum of P_{O_2} occurs 10-15 s after the flash (see records in refs 5, 12, 13). In this short time there is little diffusion of O_2 from the bath to the tip of the electrode

* Permanent address: Department of Physiology and Pharmacology, University of Pennsylvania School of Dental Medicine, Philadelphia, Pennsylvania 19104, USA.

† Present address: Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA.

and hence the minimum P_{O_2} corresponds approximately to the end of ΔQ_{O_2} . However, because the geometry of the preparation is very regular, and because Q_{O_2} is independent of P_{O_2} down to low values¹⁴, the diffusion equations relating P_{O_2} to Q_{O_2} are straightforward, and the amplitude and time course of ΔQ_{O_2} can be accurately calculated⁵ (noisy line in Fig. 1d). Because there are many mitochondria in the photoreceptors, and very few in the other cells (glia) in the centre of the retina where P_{O_2} is measured¹⁵, we assume that ΔQ_{O_2} is essentially limited to the former. These measurements of ΔQ_{O_2} and ΔNa_i^+ (supported by the measurement of ΔK_i^+) enable us to estimate the fraction of the mitochondrial respiration that is devoted to producing ATP sufficient to pump the ΔNa_i^+ out of the cell.

In response to a light flash of the same energy as that used for the Na^+ and K^+ response, the integrated extra Q_{O_2} was $3.8 \mu\text{l}$ per g of retina. If the substrate for mitochondrial respiration is pyruvate, each atom of oxygen consumed corresponds to the production of three molecules of ATP (see, for example, ref. 16), a number that is little affected by the choice of substrate. Hence, as shown in Table 1, we calculate that the mitochondria could produce ATP equivalent to a concentration of 2.0 mM in the photoreceptor. Generally, Na^+ pumps move 3 Na^+ ions for each molecule of ATP, so the oxygen measurements suggest that if all the ATP were used for pumping Na^+ , about 6.0 mM could be pumped. From the values in Table 1, we conclude that at least half, and probably more, of the energy produced by respiration is needed for the Na^+ pump.

In frog muscle, Mahler¹⁷ measured the time course of Q_{O_2} after a twitch and, finding that it lagged behind the use of ATP (primarily by the contractile apparatus), defended the assumption that an increase in ADP is the primary signal that leads to an increase in Q_{O_2} . Earlier measurements on vertebrate nerve also support this idea². In drone photoreceptors, examination of the time courses of oxygen consumption and pumping rate shows that this is not the case. After a light flash, ΔQ_{O_2} reaches its peak in 2–4 s and falls to 10% within 14 s (Fig. 1d). In contrast, Na_i^+ falls more slowly. Figure 1b shows the fastest recovery recorded in over 20 observations. Using the electrode calibration curve, the Na^+ voltage signal in Fig. 1b was converted into values of ΔNa_i^+ at a series of time intervals after the flash and these values are plotted as filled circles in Fig. 1d

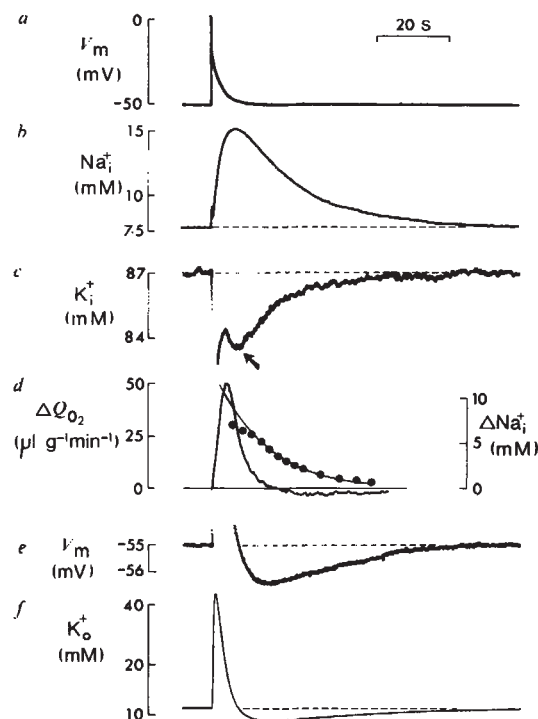


Fig. 1 Comparison of electrical, ionic and oxygen responses to single light flashes. Slices of drone retina 300–800 μm thick were stimulated with a light flash 20–40 ms long. *a*, Chart recording of an intracellular receptor potential. *b*, Sodium signal recorded as the difference in potential between barrels of a double-barrelled intracellular electrode. *c*, Potassium signal recorded with a double-barrelled intracellular electrode. *d*, Noisy line is ΔQ_{O_2} . Filled circles show ΔNa_i^+ measured from the record in *b*. The smooth curve is an exponential with $\tau = 14$ s fitted to the later points. *e*, Receptor potential recorded at a higher gain than in *a* to illustrate the afterhyperpolarization. *f*, Change in K_o^+ measured with a double-barrelled K^+ -sensitive electrode. Note the similarity in time course of the recoveries in *b*, *c*, *e* and *f*. The Na^+ sensor used in *b* is described in ref. 35; the reference barrel contained 1 M KCl. This record was selected because it has the fastest recovery observed; it was obtained from a cell that had been exposed to low Na^+ Ringer (2 mM) for 18 min, 34 min before the illustrated recording. (Such treatment consistently lowered Na_i^+ and increased ΔNa_i^+ in response to a flash¹¹.) The K^+ sensor used in *c* contained 3% K tetra-*p*-chlorophenylborate in 2,6-dimethylnitrobenzene; the reference barrel contained 1 M LiAc. The peak of the response (arrow) can be distinguished from a preceding electrical artefact. The artefact arises primarily because the sensor barrel of the electrode has a high resistance and the receptor potential is not faithfully recorded: when the signal from the reference barrel is subtracted a negative-going difference remains. The form of the artefact, but not the response, could be altered by varying the capacity compensation of the two barrels. The voltage amplitude of the K^+ signal was only 0.87 mV compared with 11.5 mV for the Na^+ signal in *b*. The response illustrated had the fastest recovery of the 12 cells observed. Record *a* was made with the reference barrel of such a double-barrelled K^+ -sensitive electrode. In *d*, ΔQ_{O_2} was computed from simultaneous measurements of P_{O_2} for 50 s made with two O_2 microelectrodes, one at the surface of the slice and the other 200 μm deep⁵. The computation depends on measured values for O_2 solubility α and diffusion coefficient D (ref. 14). An error in the value of α could affect the amplitude but not the time course; computations using other values for D showed that no plausible error in D could significantly modify the time course. The average baseline O_2 consumption was $18 \mu\text{l g}^{-1} \text{min}^{-1}$. The time for ΔQ_{O_2} to fall from its peak to $1/e$ was 4.5 s. In *e*, to reduce the prolonged depolarization seen with strong light³⁶, the flash energy was reduced to one-quarter. Such a reduction does not affect the kinetics of ΔNa_i^+ or ΔQ_{O_2} (refs 5, 11). The Ringer solution contained in mM: NaCl 270, KCl 10, $CaCl_2$ 1.6, $MgCl_2$ 10, Tris buffer 10, pH 7.4. Temperature 21–23 °C. The light source was a 150 W xenon arc; the flash delivered to the preparation, 7×10^8 photons μm^{-2} , as measured with a silicon photodiode.

Table 1 Respiration and ion fluxes

Factor	Values
Energy of flash	3.4×10^8 photons μm^{-2} (xenon lamp, glass optics)
Integrated extra O_2 consumption ($\int \Delta Q_{O_2} dt$)	$3.8 \mu\text{l}$ per g retina s.e. = $0.2 \mu\text{l}$ per g, $n = 10$
Calculated ATP production	2.0 mmol per l photoreceptor
Increase in sodium (ΔNa_i^+)	3.6 mM s.e. = 0.3 mM, $n = 8$
Decrease in potassium (ΔK_i^+)	3.6 mM s.e. = 0.5 mM, $n = 6$
Calculated ATP to pump out ΔNa_i^+	1.2 mmol per l photoreceptor

The table compares the quantity of ATP that could be produced by the extra oxygen consumed after a flash with that required to restore the ionic gradients. The value given for $\Delta Q_{O_2} dt$ is the integral of ΔQ_{O_2} from the time of the flash until ΔQ_{O_2} falls to zero. Q_{O_2} is obtained per unit volume of retina: to estimate the Q_{O_2} per unit volume of photoreceptor we have assumed that in the middle of the retina, where the P_{O_2} measurements were made, the photoreceptors occupy half the total volume¹⁵. The values for ΔNa_i^+ and ΔK_i^+ are probably underestimates because in cells with low resting potentials they were smaller or even reversed; hence, it is likely that in unimpaled cells ΔNa_i^+ and ΔK_i^+ are larger. The table suggests that the ratio of the ATP needed for the pump to the total available is at least $1.2/2.0 = 0.6$.

Table 2 Time from flash to 90% recovery for O_2 consumption and indices of Na^+ pumping

Parameter	Time
ΔQ_{O_2}	11.2 s, range 9.8–13.3 s $n = 10$
ΔNa_i^+	>32 s
ΔK_i^+	>35 s
Photoreceptor afterhyperpolarization	47 s, range 26–78 s
Undershoot of K_0^+	70 s, range 50–120 s

For ΔNa_i^+ and ΔK_i^+ the shortest times observed are given. Further, we assumed that the true maximum change was instantaneous and therefore extrapolated the exponential phase of the recovery to zero time. For the afterhyperpolarizations in this sample, the stimulus energy was reduced to 2×10^7 photons μm^{-2} .

for comparison with the ΔQ_{O_2} record: oxygen consumption has fallen to its baseline value while Na^+ pumping continues. To make the comparison even more conservative, we considered the possibility that the peak of ΔNa_i^+ was artefactually delayed and decreased in amplitude. Only part of the delay is due to the response time of the electrode itself. A possible explanation for the rest of the delay is that Na^+ diffuses within the cell to reach the electrode because local damage at the site of impalement reduces the light sensitivity there¹⁸. We therefore considered the idea that the true change in Na_i^+ is an instantaneous increase followed by an exponential decline. Following the first 10 s after the flash, the falling phase of the observed Na_i^+ was indeed exponential and the corresponding curve has been drawn in on Fig. 1d. Even with this extreme assumption, it takes 14 s for ΔNa_i^+ to fall to $1/e$ of its supposed peak value. This is still much slower than the decline of ΔQ_{O_2} where the comparable time constant is less than 5 s (ref. 5). Because this seemed a surprising result, we looked for supporting evidence.

Figure 1c is a recording of ΔK_i^+ . As the resting K_i^+ is high, the fractional change is small and the beginning of the response is marred by an electrical artefact that is explained in Fig. 1 legend. However, measurements on 12 cells consistently showed a time course of recovery indistinguishable from that of ΔNa_i^+ . Another quantity, one that can be measured with fine, single-barrelled microelectrodes, is the hyperpolarization normally observed after a receptor potential (Fig. 1e). The hyperpolarization is abolished by inhibitors of the Na^+ pump, such as strophanthidin (F. Baumann, personal communication). We propose that it reflects the activity of the Na^+ pump, either directly, because the pump is electrogenic, or indirectly, because pumping tends to reduce K^+ concentration in the extracellular space (K_0^+) (see below). An electrogenic Na^+ pump has been reported in other invertebrate photoreceptors^{19–21}. This afterhyperpolarization took at least as long to return to the baseline as did ΔNa_i^+ and ΔK_i^+ . Finally, we show (Fig. 1f) a quantity, ΔK_0^+ , that is measured, like the P_{O_2} , with an extracellular microelectrode. After a flash, K_0^+ rises to a peak that rapidly declines, presumably because K^+ enters glial cells^{10,13}. It then undershoots the baseline. This undershoot is abolished by strophanthidin (M.T. and J.A.C., unpublished) and so is probably due, in part at least, to the pumping of K^+ into the photoreceptors (see also refs 22, 23). The undershoot in Fig. 1f takes 40 s for 90% recovery from the peak, which is in agreement with the other indices of Na^+ pumping. Hence, as summarized in Table 2, four different kinds of measurement all support the idea that the time course of Na^+ pumping is slower than that of extra O_2 consumption.

We conclude that the observed extra O_2 consumption cannot be controlled in any simple way by the changes in ADP, ATP and P_i caused by the Na^+ pump. We must, therefore, look for other phenomena with time courses that precede that of ΔQ_{O_2} and that might be involved in the stimulation of the respiration.

One possibility is an increase in cytosolic Ca^{2+} , which is known to rise transiently after a light flash in invertebrate photoreceptors^{24–26}. Ca^{2+} at physiological concentrations can stimulate mitochondrial respiration apparently either by decreasing the electrical potential across the inner mitochondrial membrane (see, for example, ref. 27) or by stimulating mitochondrial dehydrogenases²⁸, although Ca^{2+} sensitivity has not been found in dehydrogenases of insects²⁹. However, if the ΔQ_{O_2} in drone retina is stimulated by the increase in cytosolic Ca^{2+} , it is unlikely to be in a straightforward way: removal of Ca^{2+} from the bathing solution, which apparently reduces free intracellular Ca^{2+} (ref. 26), increases, rather than decreases, ΔQ_{O_2} (ref. 5). Ca^{2+} can also stimulate glycogenolysis, and this is thought to lead to a contribution to the increase in respiration following stimulation of mammalian nerve and "provide in advance for the ATP that will soon be needed by the sodium pump" (ref. 30). However, in drone retina the glycogen seems to be located not in the photoreceptors but in the glial cells¹⁵. Hence, although a role for Ca^{2+} cannot be excluded, it seems worthwhile to consider alternative hypotheses. Stimulation of the photoreceptors causes an increase in the turnover of the glycogen in the glia³¹, and presumably a metabolic substrate is supplied to the photoreceptors. Conceivably, oxygen uptake by the mitochondria in the photoreceptors is substrate-limited and is activated by the arrival of this substrate from the glia. A substance that is released from the photoreceptors after a flash, and which, therefore, is a candidate for signalling to the glia, is K^+ , as previously suggested by Kuffler³². As shown in Fig. 1f, the time course of the increase in K_0^+ slightly precedes that of ΔQ_{O_2} , and it is known in other systems that raised K_0^+ can stimulate glial metabolism^{33,34}. We note, finally, that the mitochondria in the drone photoreceptors are located close to the cell membranes that face the glial cells.

We thank Messrs J.-L. Munoz, P. Perrotet and D. Pollet for technical assistance, and Professors B. Chance, J. M. Ritchie and R. W. Straub for helpful comments on the manuscript. This work was supported by Swiss NSF grant 3.399.078, NIH grant EY03504, NRSA fellowship NS 06898 (R.K.O.) and the Georges Kernen Foundation.

Received 25 October; accepted 8 December 1982.

- Chance, B. & Williams, G. R. *J. biol. Chem.* **217**, 409–427 (1955).
- Ritchie, J. M. *Prog. Biophys. molec. Biol.* **26**, 147–187 (1973).
- Erecinska, M., Stubbs, M., Miyata, Y., Dittre, C. M. & Wilson, D. F. *Biochim. biophys. Acta* **462**, 20–35 (1977).
- Fulpius, B. & Baumann, F. *J. gen. Physiol.* **53**, 541–561 (1969).
- Tsacopoulos, M. & Poitry, S. *J. gen. Physiol.* **80**, 19–55 (1982).
- Colts, J. A. & Orkand, R. K. *J. Physiol., Lond.* (in press).
- Naka, K. & Eguchi, E. *J. gen. Physiol.* **45**, 663–680 (1962).
- Baumann, F. *J. gen. Physiol.* **52**, 855–875 (1968).
- Coles, J. A. & Tsacopoulos, M. *J. Physiol., Lond.* **270**, 13P–14P (1977).
- Coles, J. A. & Tsacopoulos, M. *J. Physiol., Lond.* **290**, 525–549 (1979).
- Coles, J. A. & Orkand, R. K. *J. Physiol., Lond.* **332**, 16P–17P (1982).
- Tsacopoulos, M. & Lehmenkühler, A. *Experientia* **33**, 1337–1338 (1977).
- Coles, J. A. & Tsacopoulos, M. *J. exp. Biol.* **95**, 75–92 (1981).
- Tsacopoulos, M., Poitry, S. & Borsellino, A. *J. gen. Physiol.* **77**, 601–628 (1981).
- Perrelet, A. Z. *Zellforsch. mikrosk. Anat.* **108**, 530–562 (1970).
- Stryer, L. in *Biochemistry*, 344 (Freeman, San Francisco, 1975).
- Mahler, M. *J. gen. Physiol.* **71**, 559–580 (1978).
- Carreras, J. thesis, Geneva Univ. (1978).
- Brown, J. E. & Lisman, J. E. *J. gen. Physiol.* **59**, 720–733 (1972).
- Koike, H., Brown, H. M. & Hagiwara, S. *J. gen. Physiol.* **57**, 723–737 (1971).
- Edgington, D. R. & Stuart, A. E. *J. gen. Physiol.* **77**, 629–646 (1981).
- Heinemann, U. & Lux, H. D. *Brain Res.* **93**, 63–76 (1975).
- Kriz, N., Sykova, E. & Vyklícký, L. *J. Physiol., Lond.* **249**, 167–182 (1975).
- Brown, J. E. & Blinks, J. R. *J. gen. Physiol.* **64**, 643–665 (1974).
- Bader, C. R., Baumann, F. & Bertrand, D. *J. gen. Physiol.* **67**, 475–491 (1976).
- Levy, S. & Fein, A. *Invest. Ophthalm. Suppl.* **22**, 80 (1982).
- Akerman, K. E. O. & Nicholls, D. G. *Rev. Physiol. Biochem. Pharmac.* **95**, 149–201 (1982).
- Denton, R. M. & McCormack, J. G. *Clin. Sci.* **61**, 135–140 (1981).
- McCormack, J. G. & Denton, R. M. *Biochem. J.* **196**, 619–624 (1981).
- Landowne, D. & Ritchie, J. M. *J. Physiol., Lond.* **212**, 503–517 (1971).
- Evêquoz, V., Stadelmann, A. & Tsacopoulos, M. *J. comp. Physiol.* (in press).
- Kuffler, S. W. *Proc. R. Soc. B168*, 1–21 (1967).
- Salem, R. D., Hammerslag, R., Bracho, H. & Orkand, R. K. *Brain Res.* **86**, 499–503 (1975).
- Pentreath, V. W. & Kai-Kai, M. A. *Nature* **295**, 59–61 (1982).
- Steiner, R. A., Oehme, M., Ammann, D. & Simon, W. *Analyt. Chem.* **51**, 351–353 (1979).
- Baumann, F. & Hadjilazaro, B. *Vision Res.* **12**, 17–31 (1972).

Immunohistochemical detection of growth hormone-releasing factor in brain

Bertrand Bloch, Paul Brazeau, Nicholas Ling, Peter Bohlen, Frederick Esch, William B. Wehrenberg, Robert Benoit, Floyd Bloom* & Roger Guillemin

Laboratories for Neuroendocrinology and *The Arthur Vining Davis Center for Behavioral Neurobiology, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

The concept of a hypothalamic neurohumoral control for anterior pituitary secretion¹ postulates the existence of a growth hormone-releasing factor (GRF) of neuronal origin that stimulates the pituitary gland to release growth hormone (GH). Such a compound has not yet been isolated and characterized from the brain, although there is extensive physiological and biochemical evidence for its existence (reviewed in ref. 2). However, a 44-amino-acid amidated peptide having the physiological properties of GRF as well as chemical similarities was recently isolated from a human pancreatic tumour that had caused acromegaly³. Two shorter biologically active fragments of 40 and 37 residues were also isolated. The synthetic replicates of these human pancreas GRF (hpGRF) peptides specifically stimulate GH release *in vitro* and *in vivo*³. Assuming similarity or identity between the putative hypothalamic GRF and the tumour-derived hpGRF, we have used immunohistochemistry to search for hpGRF-like immunoreactivity in the brain. We report here that antisera against the hpGRF₁₋₄₀ peptide specifically stain neuronal cell bodies in the arcuate nucleus of the primate hypothalamus, with fibres projecting to the median eminence and ending in contact with portal vessels. This topography is characteristic of a neuronal system elaborating a releasing factor. These results provide evidence that hypothalamic GRF is very similar, if not identical, to hpGRF.

Antisera were prepared by immunizing rabbits with a mixture of synthetic hpGRF₁₋₄₀ methylated bovine serum albumin and Freund's complete adjuvant, according to a previously described procedure⁴; bleeds L455₂ and L455₃ that gave the best stainings were predominantly used for the present study. Tissues studied were brains from five normal adult rats, hypothalamus from one ox, hypothalami from squirrel monkeys which were either untreated (three animals) or injected with 1.5 mg or 100 µg colchicine into the lateral ventricle (two animals) 7 and 24 h before death, and basal hypothalami from two human adults obtained during autopsy. Tissue samples were fixed by perfusion (rats and three monkeys) or by immersion (human, ox and two monkeys) in freshly prepared 4% formaldehyde in 0.12 M phosphate buffer, pH 7.2. A fragment of the human pancreatic GRF-containing tumour was also similarly processed as a control. After fixation, slabs of tissues were rinsed in phosphate buffer with increasing concentrations of sucrose (up to 18% by weight), and 10–16 µm-thick cryostat sections were prepared, mounted on gelatinized slides and stored frozen until use.

Immunohistochemical studies were performed as follows: 10 mM phosphate buffer, pH 7.2, containing 0.3% Triton X-100 was used as washing and diluting medium. Antiserum, diluted 1:100 to 1:800, was applied to sections for 16 h, then the tissue was incubated with anti-rabbit IgG linked to fluorescein (1:20 dilution for 30 min) or peroxidase (1:100 dilution for 1 h). Peroxidase activity was revealed using diaminobenzidine tetrahydrochloride and hydrogen peroxide as substrate⁵. Antisera against luteinizing hormone-releasing factor (LRF), somatostatin and corticotropin were also used for comparative topographical studies. Controls demonstrating specificity consisted of comparison of stainings in adjacent sections treated with anti-hpGRF antiserum alone, or anti-hpGRF antiserum

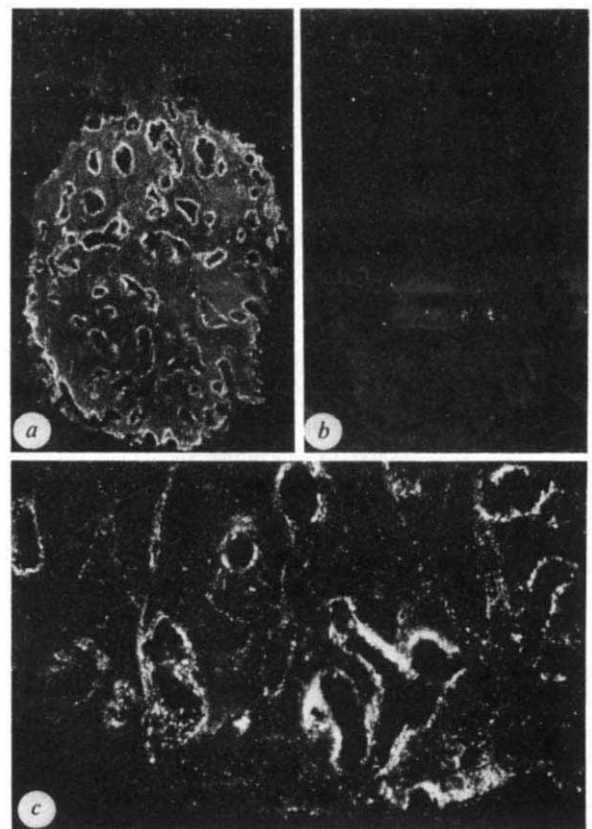


Fig. 1 Immunofluorescence staining with anti-hpGRF antiserum of median eminence of an untreated squirrel monkey. *a*, General view of the frontal section of posterior median eminence showing most of the portal vessels (dark spaces) surrounded by abundant bright immunoreactive endings. $\times 60$. *b*, The adjacent section treated with anti-hpGRF antiserum preincubated with hpGRF₁₋₄₀; no staining is apparent in the median eminence. The weakly fluorescent dots visible in the hypothalamus (upper part of the photographs) and in the median eminence correspond to autofluorescent orange pigments. $\times 60$. *c*, Higher magnification ($\times 180$) of a section stained with anti-hpGRF antiserum showing fibres terminating in strongly immunoreactive endings close to the portal capillaries.

previously incubated with hpGRF₁₋₄₀ or with synthetic fragments of the three tumoral peptides, at different concentrations.

Only primate brains revealed neuronal structures immunoreactive with the anti-hpGRF antisera used here. In human and monkey hypothalami a large set of immunoreactive fibres was found in the median eminence. In monkeys, the fibres were organized in two anterolateral bundles that spread out in the posterior part of the median eminence (Fig. 1). In both species these fibres terminate in densely packed and strongly immunoreactive endings in close contact with the portal vessels (Fig. 1). Several regions of the hypothalamus also contained scattered immunoreactive fibres. Stained cell bodies were found only in the infundibular (arcuate) nucleus; these were weakly immunoreactive and few in number in untreated monkeys, more abundant in colchicine-treated animals (up to five per section). In contrast, human brains contained numerous intensely stained cell bodies (up to several dozen per section) (Fig. 2). Neuronal structures reactive with the anti-hpGRF antisera were not stained with antiserum against somatostatin, LRF or corticotropin, each of which gave its own expected and distinct pattern of immunoreactivity in the brain. Cells of the tumour from which hpGRF was isolated also exhibited an intense cytoplasmic immunoreactivity.

In control experiments, antisera preincubated with synthetic hpGRF₁₋₄₀-OH, hpGRF₁₋₄₀-NH₂, hpGRF₁₋₄₄-NH₂ or hpGRF₂₈₋₄₄-NH₂ (0.5 mg peptide per ml of undiluted antiserum) no longer revealed immunoreactive neurones in either

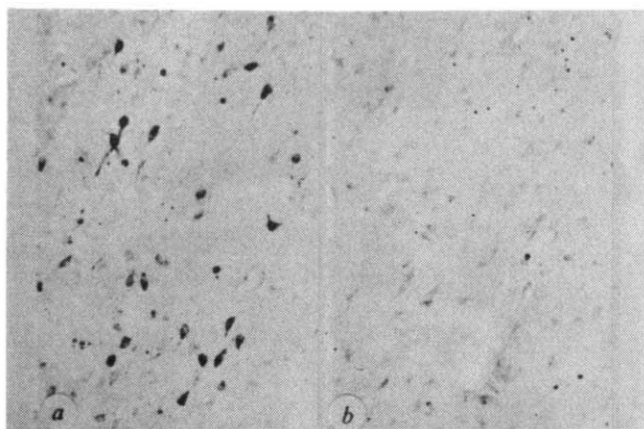


Fig. 2 Immunoperoxidase staining of the infundibular (arcuate) nucleus of a 65-yr-old man. *a*, Section stained with anti-hpGRF antiserum showing immunoreactive cell bodies and processes. $\times 50$. *b*, Absence of staining in the control adjacent section treated with anti-hpGRF antiserum preincubated with hpGRF₁₋₄₀. $\times 50$. The dark dots in *b* correspond to lysosomes or pigments containing occasional endogenous peroxidase activity.

monkey or human tissue (Figs 1, 2). In contrast, there was no diminution of the staining when the antisera were incubated with hpGRF₁₋₂₄-OH, hpGRF₁₋₃₇-OH or hpGRF₂₈₋₃₇-OH (up to 12 mg of peptide per ml of undiluted antiserum). These results indicate that the anti-hpGRF antibodies recognize the C-terminal part of the molecule in the fragment comprising residues 28 to 38 or 39.

The concept of neurohumoral control of pituitary secretion proposed by Harris¹ stipulates the presence of releasing factors or release-inhibiting factors in hypothalamic neurones whose fibres project into the median eminence and end in contact with the portal vessels. Immunohistochemical studies with antibodies against three releasing factors and one release-inhibiting factor, all isolated from the hypothalamus, namely gonadotropin-releasing factor^{6,7}, thyrotropin-releasing factor⁸, corticotropin-releasing factor⁹⁻¹² and somatostatin^{13,14} have all demonstrated such a characteristic neuronal distribution. A similar pattern of immunoreactivity found with anti-hpGRF antisera in the present study, together with the potent activity of this peptide in the specific stimulation of GH secretion³, strongly supports the hypothesis that the as yet uncharacterized hypothalamic GRF is closely related, if not identical, to the tumour-derived GRF. This hypothesis is consistent with the fact that biologically active peptides produced by tumours have always been found to be identical to the products of the physiological source of such compounds¹⁵. Also, the 28-40 region of the molecule which lacks GRF bioactivity³ probably contains structural variations among species, as the antibodies used in the present study showed no staining of neurones in ox and rat brain. However, further biochemical and immunological studies are needed to elucidate the structure of brain GRF.

The presence of immunoreactive cell bodies in the arcuate nucleus shows that this area is a primary source of GRF, at least in primates. This result agrees with recent experiments showing that selective chemical destruction of the arcuate nucleus in rats^{16,17} impairs GH secretion, presumably by destroying the source of GRF. However, the absence of stained cells in the ventromedial nucleus of primates does not correlate with the results in rats of electrical stimulation and lesion experiments that affected GH secretion and which were interpreted as proposing that this nucleus was a centre of GRF secretion^{18,19}. As our results were obtained in primates and were based on only limited experiments, involvement of the ventromedial nucleus as a source of GRF or as an activator of the arcuate GRF-producing neurones cannot be definitively excluded.

This research is supported by the NIH (grants HD-09690 and AM-18811) and the Robert J. Kleberg Jr and Helen C.

Kleberg Foundation. Bertrand Bloch is on leave from the Laboratoire d'Histologie-Embryologie, Faculté de Médecine, Besançon, France, with fellowships from the Fondation pour la Recherche Médicale and the Philippe Foundation. We thank Drs S. Foote, J. Morrison and C. Franklin for assistance in monkey tissue preparation, Drs K. Jones and P. Wolf for obtaining the human brains and B. Gayer for secretarial assistance.

Note added in proof: Additional studies in macaques and squirrel monkeys have revealed hpGRF-immunoreactive neurones in the ventromedial nucleus, although in fewer numbers than in the arcuate nucleus.

Received 19 October; accepted 9 December 1982.

1. Harris, G. *Neural Control of the Pituitary Gland* (Arnold, London, 1955).
2. Martin, J. B. in *Frontiers in Neuroendocrinology* (eds Martini, L. & Ganong, W. F.) 129-168 (Raven, New York, 1976).
3. Guillemin, R. et al. *Science* **218**, 585-587 (1982).
4. Benoit, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **79**, 917-921 (1982).
5. Sternberger, L. A. *Immunocytochemistry* (Prentice-Hall, New Jersey, 1974).
6. Leonardelli, J., Barry, J. & Dubois, M. P. *C.r. hebdom. Séanc. Acad. Sci., Paris* **D276**, 2043-2046 (1973).
7. Barry, J., Dubois, M. P. & Poulain, P. Z. *Zellforsch. mikrosk. Anat.* **146**, 351-366 (1973).
8. Hökfelt, T., Fuxe, F., Johansson, O., Jeffcoate, S. & White, N. *Eur. J. Pharmac.* **34**, 389-392 (1975).
9. Paull, W. K. et al. *Peptides* **1**, 183-191 (1982).
10. Bloom, F. E., Battenberg, E. F., Rivier, J. & Vale, W. *Regul. Peptides* **4**, 43-48 (1982).
11. Tramu, G. & Pillez, A. *C.r. hebdom. Séanc. Acad. Sci., Paris* **D294**, 107-114 (1982).
12. Bugnon, C., Fellmann, D., Gouget, A. & Cardot, J. *Nature* **298**, 159-161 (1982).
13. Dubois, M. P., Barry, J. & Leonardelli, J. *C.r. hebdom. Séanc. Acad. Sci., Paris* **D179**, 1899-1902 (1974).
14. Hökfelt, T. et al. *Acta endocr., Copenh. Suppl.* **200**, 1-41 (1975).
15. Imura, H. *Clin. Endocr. Metab.* **2**, 235-260 (1980).
16. Millard, W., Martin, J. B. Jr, Audet, J., Sagar, S. & Martin, J. B. *Endocrinology* **110**, 540-550 (1982).
17. Acs, Z., Antoni, F. & Makara, G. *J. Endocr.* **93**, 239-245 (1982).
18. Frohman, L. & Bernardis, L. *Endocrinology* **82**, 1125-1132 (1968).
19. Frohman, L., Bernardis, L. & Kant, K. *Science* **162**, 580-582 (1968).

Vasopressin alters female sexual behaviour by acting on the brain independently of alterations in blood pressure

P. Södersten*, M. Henning†, P. Melin‡ & S. Ludin‡

* Department of Psychiatry, Karolinska Institute, S-141 86 Huddinge, Sweden

† Department of Pharmacology, University of Göteborg, PO Box 33031, S-400 33 Göteborg, Sweden

‡ Ferring AB, Fack, S-200 60 Malmö, Sweden

Subcutaneous (s.c.) administration of Arg-vasopressin (AVP) prolongs retention of a learned behaviour and elevates arterial blood pressure^{1,2}. Intracerebroventricular (i.c.v.) injection of a thousandfold lower dose of AVP than needed with s.c. injection produces the same behavioural effect, suggesting that AVP acts on the brain to control behaviour³. However, as i.c.v. injection of AVP also elevates arterial blood pressure⁴, it was suggested that AVP, and perhaps other peptides as well⁵, influences behaviour indirectly by eliciting a peripheral response, for example blood pressure changes, rather than by acting directly on the brain². The suprachiasmatic nuclei (SCN) of the hypothalamus, a source of vasopressin production⁶, inhibit sexual receptivity in oestradiol-17 β -treated ovariectomized rats during the light phase of the daily lighting cycle⁷, leading to speculation that vasopressin might inhibit sexual behaviour⁸. Here we report that i.c.v. injections of AVP (1, 4 or 10 ng) inhibit sexual behaviour in receptive rats. This behavioural response is prevented by i.c.v. injection of an antiserum to AVP 30 min before AVP injection. Subcutaneous injection of a high dose of AVP (1 μ g) has no behavioural effect but elevates arterial blood pressure within 30 min of administration. Intracerebroventricular injection of a behaviourally effective dose of AVP (1 ng) has no effect on blood pressure. The results provide direct evidence that AVP can alter behaviour by an action on the brain and independently of its effect on blood pressure.

Female Wistar rats (Møllegaard Breeding Laboratories, Denmark, 200–220 g) were ovariectomized and 1 week later polyethylene (Portex Tubing PP25) or stainless-steel (Plastic Products) guide cannulae were implanted under methohexital (Brietal, Lilly) anaesthesia, in the right lateral cerebral ventricle (2 mm posterior and 2 mm lateral to the bregma and 4 mm below the surface of the skull) through burr holes and fixed to the skull with screws and dental cement. Intact male rats of the same strain (300–350 g) were cannulated with stainless-steel guide cannulae.

The following day all females were injected s.c. with 2 µg oestradiol-17β benzoate (OB, Sigma, dissolved in 0.1 ml sesame oil) at 09.00 h, 2 h before lights off in the 12-h light/12-h darkness lighting cycle and 26 h later the rats received a s.c. injection of 0.5 mg progesterone (Sigma, dissolved in 0.1 ml sesame oil). An AVP antibody (RIA titre: 1:130,000, cross-reactivity: oxytocin <0.01%, Lys-vasopressin: 19% and Arg-vasotocin: 20%) developed at Ferring AB (S.L.) was diluted in 0.9% NaCl (1:100) and injected i.c.v. in a volume of 2 µl 30 min before the AVP injection. Male rats received no hormone treatment before the AVP injection.

Tests for female sexual behaviour (lordosis behaviour) were conducted 15 min and 1 and 3 h after the AVP injection. The rats were placed individually with a sexually vigorous male rat in a 50-cm-wide circular Plexiglass arena with a sawdust-covered floor, and the animals were observed until the female had been mounted 10 times. The number of lordosis responses (concave back flexion, lateral tail deviation and neck extension) was recorded and a lordosis quotient was calculated as: (number of lordosis responses) × 10.

Tests for male sexual behaviour were conducted in the same cages as those used for studies on lordosis behaviour and were started 15 min after the AVP injection. Each male was allowed a brief adaptation period and then a sexually receptive female rat (previously injected with 20 µg OB and 0.5 mg progesterone) was introduced to the male. The number of mounts without and with intromission and the latency to the first intromission (intromission latency), the interval between the first intromission and ejaculation (ejaculation latency) and the interval between ejaculation and the next intromission (post-ejaculatory interval) were recorded.

Mean arterial blood pressure was recorded in ovariectomized rats, treated with OB and progesterone in the same way as the animals used in behavioural tests, via indwelling polyethylene catheters (Clay Adams PE50) inserted in the aorta via the left carotid artery under pentobarbital anaesthesia 48 h before measurements^{9,10}. The animals were implanted with ventricular guide cannula at the same time. Recordings were made in anaesthetized (chloral hydrate) or conscious, unrestrained animals by connecting the aortic catheter via heparin-saline-filled tubing to Statham P23 pressure transducers writing on Grass Model 7 polygraphs.

Table 1 Effect of AVP on arterial blood pressure

	Time after injection (min)				Maximum increase
	0	15	30	60	
Subcutaneous					
1 µg AVP (10)	115±6	126±7	128±6*	117±4	23±4*
NaCl (8)	104±5	108±4	110±5	111±6	8±3
Intracerebroventricular					
1 ng AVP (6)	104±4	103±4			
NaCl (7)	98±4	99±4			
25 ng AVP (8)	83±4*	82±8	76±9	80±4	

Mean ± s.e.m. arterial blood pressure (mm Hg) in ovariectomized conscious rats injected with AVP s.c. or i.c.v. Animals given 25 ng AVP were anaesthetized with chloral hydrate. Subcutaneous injection with AVP, but not NaCl, produced a significant alteration in blood pressure ($P = 0.012$, two-way analysis of variance), with an increase within 30 min of administration ($P < 0.05$, Mann-Whitney U test). The maximum increase in blood pressure after the s.c. AVP injection is indicated to the right. Numbers in parentheses refer to the number of animals per group.

* Significantly different from NaCl-treated rats, $P < 0.05$ Mann-Whitney U test.

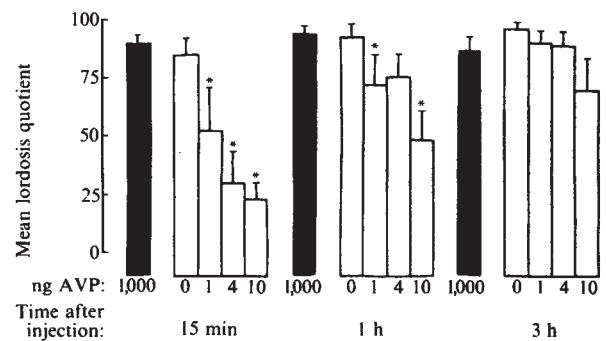


Fig. 1 Effect of intracerebroventricular (open bars) or subcutaneous (black bars) injection of AVP on sexual receptivity (mean ± s.e.m. lordosis quotient) induced by 2 µg oestradiol-17β benzoate (32 h) and 0.5 mg progesterone (3 h) before the AVP injection. AVP and Lys-vasopressin (Ferring AB) were first dissolved in 0.01 M HCl, further diluted in 0.9% NaCl and then injected i.c.v. through the guide cannula in a volume of 2 µl using a 10 µl Hamilton microsyringe 3–5 h after progesterone injection. Subcutaneous injections of AVP or NaCl were given in a volume of 0.5 ml. * Significantly different from saline-injected controls; $P < 0.05$ Mann-Whitney U test after one-way analysis of variance. There were six to eight rats per group.

Results on sexual receptivity for each test were analysed by one-way analysis of variance and subsequent between-group comparisons were made with the Mann-Whitney U test provided the analysis of variance indicated a significant overall effect (P at least < 0.05). Results on repeated measurement of arterial blood pressure after s.c. injection of AVP or NaCl were first analysed within the AVP- or NaCl-treated group using a two-way analysis of variance and between-group comparisons were made with the Mann-Whitney U test.

Figure 1 shows that i.c.v. injection of as little as 1 ng AVP inhibited the lordosis response within 15 min of administration, that inhibition was evident 1 h after injection, but that no behavioural effect was observed 3 h post-injection. Injection of the AVP antiserum 30 min before injection of 10 ng AVP prevented the inhibitory effect of the AVP (mean ± s.e.m. lordosis quotient: 88 ± 5 ($n = 5$) compared with 48 ± 2 ($n = 5$) for NaCl + AVP-treated rats, $P = 0.012$). Subcutaneous injection of a thousandfold higher dose of AVP (1 µg) produced no behavioural effect (Fig. 1), nor did injections of lower doses (10 ng, data not shown). Rats treated with NaCl i.c.v. or s.c. showed the same lordosis quotients and the results from these groups have been combined in Fig. 1.

The inhibitory effect of AVP was not overcome by increasing the dose of OB. Six rats treated with 20 µg OB followed by 0.5 mg progesterone and subsequently given 10 ng AVP i.c.v. showed a lordosis quotient of 23 ± 10 when tested 15 min later. Intracerebroventricular injection of 10 ng LVP had no inhibitory effect in females tested 15 min and 1 h after the injection (mean ± s.e.m. lordosis quotients: 80 ± 9 and 76 ± 10 ($n = 5$)).

Intracerebroventricular injection of 10 ng AVP had no significant effect on the behaviour shown by male rats in tests with sexually receptive female rats. AVP-treated animals ($n = 4$) showed 5.8 ± 2.4 mounts and 6.8 ± 1.8 intromission before ejaculating and had an intromission latency of 1.0 ± 0.8 min, an ejaculation latency of 5.0 ± 1.3 min and a post-ejaculatory interval of 6.6 ± 0.5 min. The corresponding values for the NaCl-treated rats ($n = 4$) were 3.0 ± 1.2 , 10.8 ± 2.8 , 0.2 ± 0.1 min, 4.9 ± 0.5 min and 5.1 ± 0.6 min.

1 µg AVP given s.c. produced a significant increase in arterial blood pressure within 30 min of injection (Table 1). Two of the AVP-treated rats had high initial values and the basal blood pressure, therefore, appeared higher in these animals. 1 ng AVP injected i.c.v. had no effect on blood pressure within 15 min of administration (Table 1).

Blood pressure recordings for up to 1 h after i.c.v. AVP injections consistently failed to affect blood pressure. An attempt was therefore made to replicate the recent findings that high doses of AVP i.c.v. can elevate blood pressure in anaesthetized rats⁴. However, 25 ng AVP i.c.v. produced no change in blood pressure within 1 h in chloral hydrate-anaesthetized rats (Table 1). The basal blood pressure in these rats was reduced as a consequence of the anaesthesia.

The possibility that the behavioural effect of both peripherally and centrally injected AVP might be mediated indirectly via alterations in blood pressure rather than directly via an action on the brain² received support from the recent observation that both modes of AVP administration resulted in blood pressure elevations and that prevention of the change in blood pressure by injection of an AVP antagonist before the AVP injection inhibited the behavioural effect^{2,4}. However, in our study inhibition of sexual behaviour was noted after i.c.v. injection of a low dose of AVP (1 ng) at a time when no change in blood pressure was found. Also, s.c. injection of a high dose of AVP (1 µg) resulted in blood pressure elevations but had no effect on behaviour. These data show that some behavioural effects of AVP can be dissociated from the effect on blood pressure. Because peptides injected into the lateral brain ventricle rapidly diffuse into the brain¹¹, the data also show that AVP can act directly on the brain to alter behaviour.

We advise caution in interpreting the physiological significance of the recent demonstration that i.c.v. AVP injections can produce blood pressure elevations⁴. Extremely high doses seem to be required and we observed no effects of a behaviourally effective dose, or a 25-fold higher dose, of AVP.

The effect of AVP on female sexual behaviour is specific, as no effect was noted with a closely related peptide, Lys-vasopressin, and as the treatment failed to affect the display of sexual behaviour by male rats. Moreover, the effect was antagonized by i.c.v. injection of an AVP antiserum before the AVP injection. Peptide antisera injected i.c.v. readily penetrated the brain¹².

Ovariectomized oestradiol-17β-treated rats show a daily rhythm in sexual receptivity which is disrupted by SCN lesions⁷. The SCN produce vasopressin⁶ and vasopressin levels vary rhythmically in the cerebrospinal fluid, but not the blood, during the light:darkness cycle, with peak levels during the light phase¹³. It is during this phase of the cycle that the SCN are maximally active metabolically¹⁴ and electrophysiologically¹⁵ and, correlatively, sexual behaviour is inhibited⁷. It is tempting to speculate, therefore, that vasopressin, produced by the SCN, rhythmically inhibits the sexual behaviour of oestradiol-17β-treated rats. Inhibition by vasopressin may be one of the factors that terminate the behaviour in physiological conditions^{7,8}.

We thank Dr B. J. Everitt for his helpful comments. This work was supported by the Swedish Council for Research in the Humanities and Social Sciences (431/81), the Funds of the Karoliska Institute and the Swedish MRC (2863).

Exposure of an antigen of chromaffin granules on cell surface during exocytosis

G. Lingg, R. Fischer-Colbrie, W. Schmidt & H. Winkler

Department of Pharmacology, University of Innsbruck, A-6020 Innsbruck, Austria

The synthesis rate of the membrane proteins of the catecholamine-storing vesicles (chromaffin granules) of the adrenal medulla is lower than that of the secretory proteins of the contents¹. Based on these results we proposed^{2,3} that after exocytosis the membranes of chromaffin granules are retrieved and are re-used for several secretion cycles (see also ref. 4). This concept of re-use of granule membranes has been further strengthened by the finding that exogenous markers which are taken up by secretory cells during stimulation can be traced to the Golgi region and to immature secretory organelles^{5,6}. However, one basic question remains: are the membranes of secretory organelles specifically and completely removed from the plasma membrane and if so, how fast is this process? By using an antiserum against a membrane glycoprotein of chromaffin granules we have now obtained quantitative data which demonstrate that during exocytosis this antigen becomes exposed on the cell surface and disappears again to a large degree within 30 min.

Rabbit antisera were raised against glycoprotein III (GP III) isolated from bovine chromaffin granules as described previously⁷. This antigen is specifically confined to chromaffin granules and its antigenic groups are present on the inner side of the granule membrane facing the interior (R.F.-C. and H.W., manuscript in preparation). Isolated chromaffin cells were resuspended in 0.5% bovine serum albumin/Krebs-Henseleit buffer⁸ and incubated for 1 h at room temperature. After sedimentation they were resuspended again in the same buffer (2×10^6 cells ml⁻¹). For stimulation, carbachol was added to a final concentration of 2.2 mM. After incubation at 37 °C for 10 min and sedimentation of these cells at 0 °C, catecholamines were determined in the supernatants of control and stimulated cells. We subtracted from these levels the amounts of catecholamines present in the supernatant of resuspended cells before incubation. During a 10-min incubation, $12.5 \pm 1.0\%$ ($\pm$ s.e., $n = 11$) of the total catecholamines were secreted in the presence of carbachol as opposed to $3.6 \pm 0.6\%$ ($n = 17$) in its absence.

To test for the appearance of the granule antigen on the cell surface, two methods were used—determination of cytotoxic effect of antiserum plus complement⁹, and microcomplement fixation. To measure the cytotoxic effect of antiserum plus complement, antiserum against GP III (1:160 dilution) and complement (1:50 dilution of guinea pig serum) were added to cells incubated for 10 min at 37 °C with (stimulation) and without (control) carbachol. After incubation these cells were chilled and Trypan blue was added (final concentration 0.6 mg ml⁻¹). After 2 min the samples were fixed overnight with 0.5% glutaraldehyde. The percentage of Trypan blue-stained cells (damaged by cytotoxic effect), determined in a haemocytometer, was $30 \pm 1.9\%$ ($n = 17$, range 21–50%) in control cells compared with $49 \pm 2.7\%$ ($n = 17$, range 32–66%) in stimulated cells. The values for the two samples obviously overlap, however, a clearcut difference between stimulated and control cells could be observed in each experiment from 17 different cell preparations. This difference (see Fig. 1), on average, was $19 \pm 1.8\%$ ($n = 17$) and as shown by the standard error was reproducible. Control sera had no cytotoxic effect.

Received 23 August; accepted 21 December 1982.

- De Wied, D. *Nature* **232**, 58–60 (1971).
- Le Moal, M. et al. *Nature* **291**, 491–493 (1981).
- De Wied, D. *Life Sci.* **19**, 685–690 (1976).
- Pittman, Q. J., Lawrence, D. & McLean, L. *Endocrinology* **110**, 1058–1060 (1982).
- Iversen, S. D. *Nature* **291**, 454 (1981).
- Sofroniew, M. V. & Weindl, A. in *Endogenous Peptides and Learning and Memory Processes* (eds Martinez, J. L. Jr, Jensen, R. A., Messing, R. B., Rigter, H. & McGaugh, J. L.) 327–369 (Academic, New York, 1981).
- Hansen, S., Södersten, P., Eneroth, P., Srebro, B. & Hole, K. *J. Endocr.* **83**, 267–274 (1979).
- Södersten, P., Eneroth, P. & Hansen, S. in *Steroid Hormone Regulation of the Brain* (eds Fuxe, K., Wetterberg, L. & Gustafsson, J.-Å.) 301–315 (Pergamon, New York, 1981).
- Popovic, V. & Popovic, P. *J. appl. Physiol.* **15**, 727–728 (1960).
- Henning, M. *Acta physiol. scand. suppl.* **332** (1969).
- Jackson, I. M. D. in *Neurobiology of Cerebrospinal Fluid* (ed. Wood, J. H.) 625–650 (Plenum, New York, 1980).
- Swaab, D. F. & Boer, K. in *Animal Models in Fetal Medicine* (ed. Nathanielsz, P. W.) 169–234 (Elsevier, Amsterdam, 1981).
- Reppert, S. M., Artman, H. G., Swaminathan, S. & Fisher, D. A. *Science* **213**, 1256–1257 (1981).
- Schwartz, W. J., Davidsen, L. C. & Smith, C. B. *J. comp. Neurol.* **189**, 157–167 (1981).
- Inouye, S.-I. T. & Kawamura, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5962–5966 (1979).

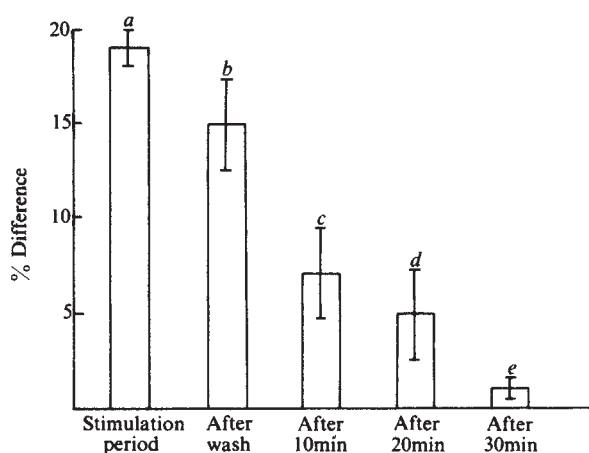


Fig. 1 Cytotoxic effect of complement plus antiserum against a granule glycoprotein on isolated chromaffin cells. Isolated cells from bovine adrenal medulla were subjected to a cytotoxicity test with antiserum and complement. The percentage of damaged cells in controls was subtracted from that of stimulated cells. These differences are shown. The time course is given on the abscissa. *a* Represents the difference in cytotoxicity between control and stimulated cells during a stimulation period for 10 min at 37 °C ($n = 17$). After this incubation the cells were washed (three times at 0 °C to remove the secretagogue carbachol) and the two samples were again tested for cytotoxicity (*b*, $n = 12$); *c-e* represent aliquots taken from the two samples reincubated for 10 min (*c*, $n = 7$), 20 min (*d*, $n = 4$) and 30 min (*e*, $n = 12$) at 37 °C. The s.e.m. is given.

From these data we suggest that during stimulation an increased amount of the granule antigen becomes exposed on the outer surface of chromaffin cells and renders the cells susceptible to the cytotoxic effect of the antiserum plus complement. Due to the complexity of the cytotoxic reaction these results are unlikely to provide a simple relationship between the amounts of antigen present on the cell surface and the degree of cytolysis observed. In any case the cytotoxicity test proved suitable for studying quantitatively the time course of the disappearance of the antigen from the cell surface after the exocytotic event. For this purpose control and stimulated cells (after 10 min at 37 °C) were sedimented, resuspended at 0 °C and incubated for up to 30 min at 37 °C. At various times aliquots were taken for the cytotoxicity test (incubation for 10 min with antiserum and complement). Figure 1 shows that the difference in cytotoxicity was still present in the resuspended cells but gradually disappeared during incubation at 37 °C and was almost completely absent after 30 min.

Preliminary studies using the cytotoxicity test indicate that analogous results could be obtained with an antiserum against dopamine β -hydroxylase. It has been reported previously¹⁰ that, when stimulated, isolated cells of bovine adrenal medulla show an increased immunohistochemical staining of this antigen on the cell surface. However, no quantitative data were given.

Microcomplement fixation was used as a second method to test for the exposure of the granule antigen on the cell surface. After incubation for 10 min (calcium-free controls and stimulated cells) the cells were fixed in 1.0% formaldehyde in 0.1 M cacodylate buffer at 0 °C for 30 min (this fixation did not decrease the antigenicity). The fixed cells were washed several times and then subjected to microcomplement fixation¹¹. One aliquot of the unfixed cells was lysed with distilled water, frozen and thawed and the total membranes were then sedimented. These membranes were used to determine the total amount of antigen present in the cells. In control preparations $5.3 \pm 0.3\%$ ($n = 3$) of the total antigen present in chromaffin cells was accessible to the antibody in microcomplement fixation. This exposed antigen may be an expression of a low exocytotic activity even in controls, but it is also possible that antigen

present in damaged cells or cell debris, might have reacted. In any case, during stimulation the percentage of accessible antigen increased to $17.5 \pm 3.7\%$ ($n = 3$). After 30 min of subsequent incubation without stimulant, this percentage was reduced to $7.5 \pm 0.2\%$ ($n = 3$). Thus, most of the antigen exposed on the cell surface by stimulation disappears again within 30 min, which is in good agreement with the first method.

Thus, using two different methods we have obtained the first quantitative data on the appearance of an antigen of secretory granules on the cell surface during exocytosis. During a 10-min stimulation about 8.9% of the catecholamines was secreted and about 12.2% of the antigen (control values subtracted) became exposed on the outer surface. Within 30 min this antigen disappeared again to a large degree. The most likely mechanism is a specific retrieval of membrane patches. By using the methods described here the effects of drugs on the retrieval mechanism can now be studied quantitatively.

This study was supported by Dr Legerlotz-Stiftung.

Received 24 August; accepted 10 December 1982.

1. Winkler, H., Schöpf, J. A. L., Hörtnagl, H. & Hörtnagl, H. *Naunyn Schmiedeberg's Archs Pharmacol.* **273**, 43–61 (1972).
2. Winkler, H., Schneider, F. H., Rufener, C., Nakane, P. K. & Hörtnagl, H. *Advances in Cytopharmacology* Vol. 2, 127–139 (Plenum, New York, 1974).
3. Winkler, H. *Neuroscience* **2**, 657–683 (1977).
4. Meldolesi, J., Borgese, N., De Camilli, P. & Ceccarelli, B. *Cell Surface Review* Vol. 5, 509–627 (North-Holland, Amsterdam, 1978).
5. Herzog, V. *Phil. Trans. R. Soc. B296*, 67–72 (1981).
6. Suchard, S. J., Corcoran, J. J., Pressman, B. C. & Rubin, R. W. *Cell-Biol. int. Rep.* **5**, 953–962 (1981).
7. Fischer-Colbrie, R., Schachinger, M., Zangerle, R. & Winkler, H. *J. Neurochem.* **38**, 725–732 (1982).
8. Fenwick, E. M., Fajdiga, P. B., Howe, N. B. S. & Livett, B. G. *J. Cell Biol.* **76**, 12–30 (1978).
9. Lee, V. H., Greene, L. A. & Shelanski, M. L. *Neuroscience* **5**, 1979–1987 (1980).
10. Wildman, J., Dewair, M. & Matthaei, H. *J. Neuroimmunol.* **1**, 353–364 (1981).
11. Schneider, F. H., Smith, A. D. & Winkler, H. *Br. J. Pharmacol. Chemother.* **31**, 94–104 (1967).

Naturally occurring anti-idiotypic antibodies in myasthenia gravis patients

Donard S. Dwyer*, Ronald J. Bradley†, C. Kendrick Urquhart† & John F. Kearney

Department of Microbiology, Cellular Immunobiology Unit and †Department of Psychiatry, University of Alabama in Birmingham, Birmingham, Alabama 35294, USA

Aberrant regulation of the immune system can lead to the development of autoimmune diseases such as myasthenia gravis. Autoantibodies against the nicotinic acetylcholine receptor (AChR) are found in the serum of myasthenia gravis patients and trigger a reduction of AChR at the muscle endplate resulting in increased muscle fatigability¹. It is possible that the autoimmune response results from altered idiotypic anti-idiotypic network interactions^{2,3}. Here we have used a monoclonal anti-AChR antibody (ACR-24, $\gamma 1$, κ) in an enzyme-linked immunosorbent assay (ELISA) to measure anti-AChR immunoglobulin in human sera. In this assay, ACR-24 is attached to microtitre plates followed by the addition of solubilized human AChR which is bound by the immobilized ACR-24. However, during the development of this assay, it was observed that certain myasthenic patients appeared to have antibodies which bound to ACR-24 alone. This unexpected finding suggested that we had discovered naturally occurring anti-idiotypic antibodies in myasthenic sera.

*Present address: Max-Planck-Gesellschaft Klinische Forschungsgruppe für Multiple Sklerose, Josef-Schneider-Straße 11, D-8700 Würzburg, FRG.

Various groups^{4,6} have now succeeded in producing monoclonal antibodies against the AChR by somatic cell hybridization, some of which have been shown to elicit experimental myasthenia gravis in rats^{7,8} or mice⁹. Recently, we have used ACR-24, a BALB/c monoclonal antibody raised against AChR from *Narcine braziliensis*, to measure anti-muscle AChR antibodies by an ELISA procedure¹⁰. The data in Table 1 were obtained from 10 myasthenia gravis patients using the ELISA and show binding of some of these sera to ACR-24 when compared with a control BALB/c $\gamma 1, \kappa$ antibody (MOPC21). Significant binding to ACR-24 was not observed in any normal controls. Normal purified human IgG was also tested and no binding to ACR-24 was observed. Thus it appeared that the binding exhibited by certain sera from myasthenia gravis patients was specific for ACR-24 and suggested that anti-idiotypic antibodies might be involved.

To examine the specificity of this binding, several myasthenia gravis patients showing the highest levels of this putative anti-idiotypic reaction were selected for further study. Serum (40–70 ml) was obtained from three patients and one normal volunteer. Immunoglobulins were precipitated with ammonium sulphate and dialysed extensively against borate-saline buffer. The immunoglobulin fraction was then passed several times over an affinity column containing purified RPC-23, an irrelevant $\gamma 1, \kappa$ immunoglobulin of BALB/c origin, in order to remove possible anti-isotype antibodies. Following this adsorption step, anti-idiotypic antibodies were purified with an ACR-24-Sepharose affinity column. After elution with glycine-HCl buffer, pH 2.8, the anti-idiotypic antibodies were concentrated and dialysed against borate-saline buffer before use in assays. Between 10 and 60 μg of protein per ml of serum were obtained from each patient by this procedure. A normal control serum processed in the same way did not yield antibody from the ACR-24 column.

Antibodies purified from the myasthenia gravis patients in this way were then tested for binding to ACR-24 as well as for human heavy- and light-chain isotype expression. Generally, the purified anti-idiotypic antibodies from serum are γ, κ class immunoglobulins, while μ and λ chains represent minor components in the material that was obtained from the ACR-24

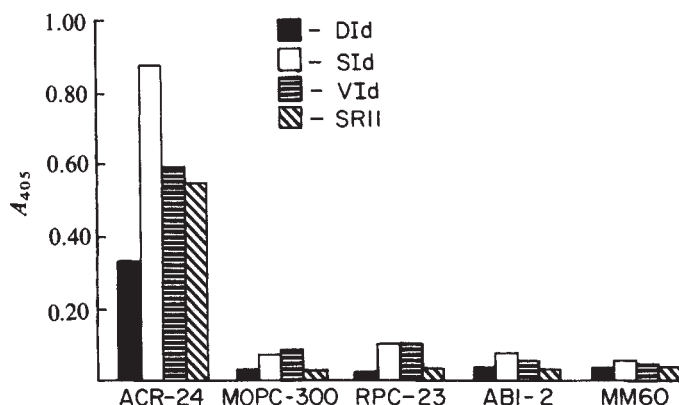


Fig. 1 Binding specificity of anti-idiotypic antibodies measured by ELISA. The results here have been obtained with anti-idiotypic antibodies purified from three myasthenia gravis patients, SId, DIId and VId. These patients showed only mild symptoms and had anti-human AChR antibody titres of 0.6, 0.9 and 1.6 nM L^{-1} as measured by the precipitation assay. In addition, data are shown for the human monoclonal antibody SR-11. Using an ELISA similar to that described in Table 1 legend, the purified anti-idiotypic antibodies were tested against: ACR-24 (BALB/c $\gamma 1, \kappa$), MOPC 300 (BALB/c $\gamma 1, \kappa$), RPC-23 (BALB/c $\gamma 1, \kappa$), ABI-2 (A/J $\gamma 1, \kappa$) and MM60 (BALB/c μ, λ). The various mouse antibodies were coated on microtitre plates at a concentration of 5 $\mu\text{g ml}^{-1}$. After blocking the wells with a 1% BSA solution, 100 μl of purified anti-idiotypic antibodies (20 $\mu\text{g ml}^{-1}$) or 100 μl of SR-11 (20 $\mu\text{g ml}^{-1}$) were added to each well. The data are expressed as A_{405} readings from the ELISA multiscan taken after 20 min of enzyme-substrate reaction.

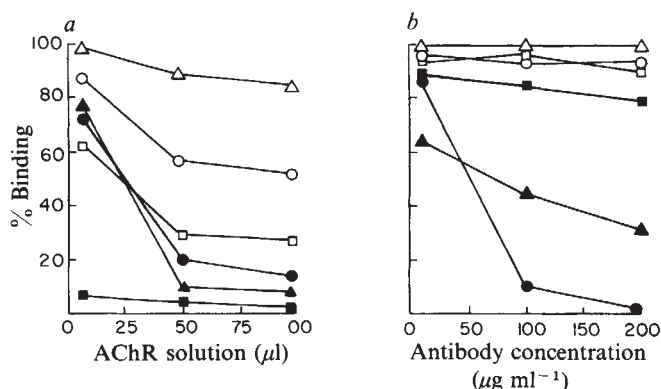


Fig. 2 Inhibition of idiotype/anti-idiotype interaction by antigen (AChR) or idiotype. *a*, The idiotype bearing antibody, ACR-24, was bound to microtitre plates followed by the addition of either a control solution, purified *Narcine* AChR, 100 $\mu\text{g ml}^{-1}$ (solid symbols), or a 2% Triton extract of denervated rat muscle containing 0.3 $\mu\text{g ml}^{-1}$ AChR protein (open symbols). Immediately after introducing these solutions, 100 μl of the purified anti-idiotypic antibodies (20 $\mu\text{g ml}^{-1}$) were added. These were SId (circles), DIId (triangles) and SR-11 (squares). Following an 8–12-h incubation period at 4 °C the plates were assayed for anti-ACR-24 activity using AP-labelled anti-human IgM or IgG antibodies. Values are expressed as % binding compared with control solutions. *b*, These assays were performed as described above except that the purified monoclonal antibodies, ACR-24 (solid symbols) and MOPC 300 (open symbols) were tested for inhibition of binding SId, DIId and SR-11 to ACR-24. The anti-idiotypic antibodies are denoted by the same symbols as in *a*.

affinity column. The purified serum antibodies were tested against a panel of murine $\gamma 1, \kappa$ monoclonal antibodies and one μ, λ to ensure that binding was specific for ACR-24 (Fig. 1). It is apparent that the purified antibodies do not react with isotypic or allotypic determinants of mouse immunoglobulins and therefore have the characteristics of anti-idiotypic antibodies. The human monoclonal antibody, SR11 (μ, κ), has recently been produced in a fusion of peripheral blood lymphocytes (PBL) from a myasthenic patient with the human myeloma line GK-5. GK-5 is a variant of GM 1500 (ref. 11) and synthesizes only a κ light chain (J.F.K. and L. Gartland, unpublished observations). It has not been established yet whether the SR 11 cell line is a true hybridoma, however, immunoglobulin-producing clones from the original line have been grown continuously in culture for over 4 months. The SR 11 monoclonal antibody has similar binding properties and specificity for ACR-24 as the purified serum antibodies from other myasthenic patients and thus may be considered a monoclonal anti-idiotypic antibody. Recently, a second monoclonal antibody with similar specificity was isolated from a fusion of GK-5 with PBL from a different myasthenia gravis patient (results not shown).

With other anti-idiotypic antibody systems, it has typically been found that antigen inhibits binding of the anti-idiotypic antibody to idiotype^{12,13}. Figure 2*a* shows the results of experiments designed to test this possibility. The addition of AChR, especially that purified from electroplax of *Narcine braziliensis*, produces marked inhibition of anti-idiotypic binding to ACR-24. The superiority of *Narcine* AChR over muscle AChR could be due to at least two factors: first, there is simply more *Narcine* receptor than muscle receptor present in the wells during the assay and second, ACR-24 has a higher affinity for *Narcine* AChR than muscle AChR. We show that relatively little muscle AChR can inhibit binding of the human monoclonal IgM (SR 11) to ACR-24. Although 100 μl of ACR-24 (5 $\mu\text{g ml}^{-1}$) were added to each ELISA well, only a portion of the total antibody actually adheres to the plastic. Assuming that approximately 1/10th of the added protein binds to the plate, and that ACR-24 has a greater affinity for AChR than it does for SR 11, then it is not surprising to observe inhibition by small amounts of muscle AChR.

We determined whether ACR-24 itself or other $\gamma 1$, κ antibodies inhibited the idiotypic interaction (Fig. 2b). It is clear that ACR-24 inhibits binding of the serum anti-idiotypes while MOPC 300 has no effects. Binding of the human monoclonal anti-idiotypic is only partially diminished by ACR-24; however, due to the greater avidity of IgM antibodies it is not unexpected that binding of SR 11 to ACR-24 is more difficult to inhibit.

One possible explanation for the above results could be that the anti-idiotypic antibodies were complexes of AChR-anti-AChR antibodies and it was these anti-receptor antibodies that were being measured in our assay. However, microtitre plates coated with other murine anti-AChR monoclonal antibodies did not bind the purified serum anti-idiotypic antibodies. Furthermore, when the purified antibodies were run on SDS gels, only two major bands corresponding to heavy and light chains were seen (data not shown). Finally, tissue culture fluid supernatant from SR 11 behaves in the same way as the purified serum anti-idiotypic antibodies; this would exclude the likelihood of AChR complex formation. Therefore, we conclude that the immunoglobulin purified with the ACR-24 affinity column is anti-idiotypic antibody. Each of the three patients discussed here had more than one subclass present in the purified serum anti-idiotypes, and in fact, each patient had a different predominant subclass—SID- $\gamma 1$, DID- $\gamma 4$ and VID- $\gamma 3$ —which suggests that the anti-idiotypic antibodies are rather heterogeneous.

To determine whether certain sera from myasthenia gravis patients contain antibodies which share the ACR-24 idiotype, we chose two patients who had demonstrable anti-idiotypic antibodies during the course of their disease for detailed analysis. From a series of serum specimens of these two particular patients, samples were selected which contained high anti-AChR titres but low anti-idiotypic levels. The sera were tested for inhibition of binding of the purified anti-idiotypic antibodies to ACR-24 as measured by the ELISA, and further, the purified serum anti-idiotypic antibodies (SID and DID) were tested for their ability to inhibit the interaction of these two sera with human AChR in an immunoprecipitation assay. The results in Fig. 3a, b show a clear inhibition of anti-idiotypic binding to ACR-24 in the presence of heterologous sera thought to contain an idiotypic counterpart to ACR-24. In addition, the purified anti-idiotypic antibodies prevented binding of some anti-AChR antibodies from serum to labelled AChR in the immunoprecipitation assay.

Thus we have demonstrated that anti-idiotypic antibodies may occur naturally during the course of myasthenia gravis. Comparable findings have been reported in the case of another human autoimmune disease, systemic lupus erythematosus¹⁴ and anti-idiotypic antibody has been measured in the serum of some IgA-deficient patients¹⁵. Shared idiotypic between mouse and human antibodies of the same specificity has also been previously reported by others^{16–18} and a monoclonal idiotypic specific antibody for a human anti-intermediate filament monoclonal antibody reacts strongly with a mouse antibody of similar specificity (A. Landay and M. D. Cooper, personal communication). These results suggest that shared idiotypic between antibodies from different species but with similar antigen binding specificities may not be uncommon.

We have begun to examine individual myasthenic sera to determine the frequency of occurrence of these anti-idiotypic antibodies. Overall, 40% of more than 50 myasthenic patients tested so far have antibodies specific for ACR-24. We have also found that the highest levels are in untreated patients with very low anti-AChR titres, while patients with high titres often have very little detectable anti-idiotypic immunoglobulin. The measurement of anti-ACR-24 antibody in patients with very low anti-AChR titres may therefore be a useful diagnostic test.

The coincidental association of high levels of anti-idiotypic with low levels of anti-AChR antibodies suggests that there may be a suppressive effect on anti-receptor antibodies or B-cell progenitors. Alternatively, the observed distribution of anti-idiotypic perhaps suggests that the anti-idiotypic is directed

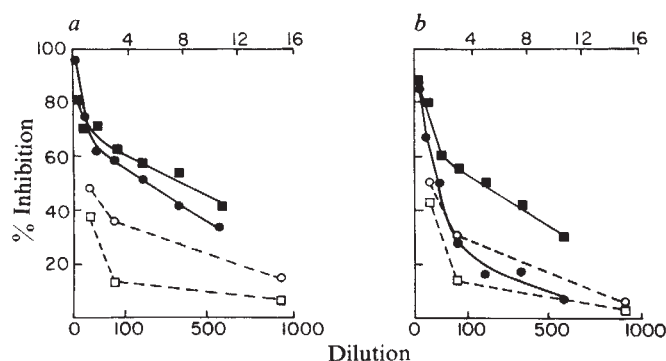


Fig. 3 Detection of ACR-24 idiotype in selected myasthenic sera with low anti-ACR-24 titre. *a*, Sera from two patients (JA, anti-AChR serum titre 890 nM; anti-idiotypic ELISA value 0.071, and AW, anti-AChR titre 21.8 nM; anti-idiotypic level 0.030) were tested for the presence of antibodies sharing the ACR-24 idiotype by two methods. First, microtitre plates were coated with ACR-24, followed by the addition of purified anti-idiotypic antibodies (SID) and increasing dilutions of the two sera. The plates were developed with goat AP-anti-human IgG and inhibition of binding to ACR-24 was determined by comparison with a normal serum diluted in the same way. These data are depicted with solid symbols: JA (—■—) and AW (—●—). The second method involved the inhibition of binding of the serum antibodies (JA and AW) to human AChR by SID. Thus, the sera were tested for immunoprecipitation of ¹²⁵I-labelled α -bungarotoxin-AChR complexes (as described elsewhere²⁰) in the presence of several dilutions of SID. Immunoprecipitation values with buffer alone were used to derive the inhibition data which are represented by open symbols and dashed lines: JA (---□---), AW (---○---). The horizontal axis at the top of this graph represents dilution of antibodies in precipitation assays. Normal human IgG were not inhibitory in this assay. *b*, Data obtained as described in *a* except that the purified anti-idiotypic antibodies, DID from a different patient, were used. The symbols used are the same as described in Fig. 3a.

Table 1 A₄₀₅ values for serum from myasthenia gravis patients binding to ACR-24 and a control and antibody

Serum no.	Bovine serum albumin	ACR-24	Control monoclonal antibody
1	0.019	0.435	0.033
2	0.028	0.347	0.045
3	0.041	0.275	0.046
4	0.003	0.062	0.020
5	0.002	0.046	0.012
6	0.018	0.198	0.047
7	0.023	0.114	0.014
8	0.019	0.252	0.010
9	0.024	0.008	0.018
10	0.012	0.044	0.000
Control*	0.011 (0.006)	0.013 (0.015)	0.009 (0.011)
IgG [†]	0.042	0.068	0.057

The ELISA was performed, as described elsewhere¹⁰, by coating microtitre plates with ACR-24 (5 μ g ml⁻¹), MOPC 21 (5 μ g ml⁻¹) or a 1% bovine serum albumin (BSA) solution. Serum was added at a 1:40 dilution in borate-saline buffer containing 1% BSA and 0.05% Tween and incubated for 2 h at room temperature. Following the addition of serum, the plates were washed and alkaline phosphatase (AP)-coupled goat anti-human IgG antibodies were added. Plates were developed with substrate (*p*-nitrophenyl phosphate) and read with an ELISA multiscan (Flow Laboratories) after 30 min.

* The control data were obtained with sera from 20 normal volunteers and are represented as the mean and s.d. of serum values measured with the ELISA.

† Normal purified human IgG (300 μ g ml⁻¹) was also tested in the same way and showed no significant binding to ACR-24.

against T-cell receptors rather than B-cell products. There is some speculation that the appearance of helper T cells sensitized to thymic AChR precedes a peripheral response in myasthenia gravis¹⁹. If a physiological role for these anti-idiotypic antibodies in myasthenia gravis can be found it will be important both in the study of idiopathic regulation of the immune response and as a possible means of specific immunotherapy in autoimmune diseases.

We thank Dr S. J. Oh for providing patients' sera, Dr H. Kubagawa for discussions and the Alabama Chapter of the Myasthenia Gravis Foundation for financial support. Purified Narcine AChR was provided by Dr G. Kemp. This work was also supported by NCI grants CA 16673 and CA 13148, and grant AI 14782, from the National Institute of Allergy and Infectious Diseases. J.F.K. is the recipient of Research Career Development Award AI 00338. D.S.D. is the recipient of a Muscular Dystrophy Association postdoctoral fellowship.

Received 6 October; accepted 7 December 1982.

1. Lindstrom, J. & Dau, P. *Ann. Rev. Pharmac. Tox.* **20**, 337 (1980).
2. Jerne, N. K. *Ann. Immun. (Inst. Pasteur, Paris)* **125C**, 373 (1973).
3. Herzenberg, L. A., Black, S. J. & Herzenberg, L. A. *Eur. J. Immun.* **10**, 1 (1980).
4. Gomez, C. M. *et al. Biochem. Biophys. Res. Commun.* **88**, 575 (1979).
5. Moshly-Rosen, D., Fuchs, S. & Eshhar, Z. *FEBS Lett.* **106**, 389 (1979).
6. Tzartos, S. J. & Lindstrom, J. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 755 (1980).
7. Lennon, V. A. & Lambert, E. H. *Nature* **285**, 238 (1980).
8. Richman, D. P. *et al. Nature* **286**, 738 (1980).
9. Dwyer, D. S., Kearney, J. F., Bradley, R. J., Kemp, G. E. & Oh, S. J. *Ann. N.Y. Acad. Sci.* **377**, 143 (1981).
10. Dwyer, D. S., Bradley, R. J., Urquhart, C. K. & Kearney, J. F. *J. Immun. Meth.* (in the press).
11. Croce, C. M., Linnenbach, A., Hall, W., Steplewski, Z. & Koprowski, H. *Nature* **288**, 488 (1980).
12. Sher, A. & Cohn, M. *J. Immun.* **109**, 176 (1972).
13. Clafin, J. L., Lieberman, R. & Davie, J. M. *J. Immun.* **112**, 1747 (1974).
14. Abdou, N. I., Wall, H., Lindsley, H. B., Halsey, J. F. & Suzuki, T. *J. clin. Invest.* **67**, 1297 (1981).
15. Cunningham-Rundles, C. *J. exp. Med.* **155**, 711 (1982).
16. Riesen, W. F. *Eur. J. Immun.* **9**, 421 (1979).
17. Bast, B. J., Copper, M. D. & Kearney, J. F. *J. Immun.* (submitted).
18. Schwartz, M., Novick, D., Givol, D. & Fuchs, S. *Nature* **273**, 543 (1978).
19. Wekerle, H., Hohfeld, R., Keteisen, U.-P., Kalden, J. R. & Kalies, I. *Ann. N.Y. Acad. Sci.* **377**, 455 (1981).
20. Dwyer, D. S., Bradley, R. J., Oh, S. J. & Kemp, G. E. *Clin. exp. Immun.* **37**, 488 (1979).

Alloantigen-specific suppressor T cells can also suppress the *in vivo* immune response to unrelated alloantigens

A. T. J. Bianchi, H. Bril & R. Benner

Department of Cell Biology and Genetics, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

Delayed-type hypersensitivity (DTH) to both major histocompatibility complex (H-2) and non-H-2-coded antigens can be induced by subcutaneous immunization with allogeneic lymphoid cells in the mouse. While subcutaneous immunization with allogeneic cells preferentially induces DTH reactivity, intravenous immunization, especially with irradiated allogeneic cells, induces a state of suppression. Suppression is manifest both in direct host-versus-graft (HvG)¹ assays and under graft-versus-host (GvH) conditions², where spleen cells of suppressed mice are used to reconstitute irradiated allogeneic hosts. The suppression is mediated by T cells^{2,3}. We have now studied the specificity of the suppressive effect by subcutaneous immunization of 'suppressed' mice with a combination of alloantigens comprising the antigen(s) used to induce the suppressor T cells as well as unrelated alloantigens. We report here that reaction against the third party alloantigens was effectively suppressed, provided these antigens were presented in combination with the antigen(s) that had induced the suppressor T cells. Both sets of alloantigens do not need to be physically associated.

Mice, suppressed to H-2 histocompatibility antigens, do not display a DTH response against these antigens, but respond normally against third party antigens (Table 1, groups A and E). The suppression is H-2 subregion-specific: mice, suppressed for either H-2I or H-2D antigens, display normal DTH reac-

tivity after subcutaneous immunization with other H-2 subregion antigens (Table 1, groups B and C). Table 1, group D, shows that suppression of the HvG reactivity to H-2 subregions is haplotype-specific. We have also observed that intravenous injection of cells incompatible for minor histocompatibility antigens could not suppress DTH reactivity to subcutaneously administered H-2 alloantigens and vice versa. Thus suppressor T cells require restimulation with the original immunizing antigen: they are strictly antigen-specific as far as their activation is concerned.

In GvH reactions H-2I and IIs antigens activate DTH-reactive T cells⁴. This activity is assayed in secondary recipients, after passive transfer of lymphoid cells from animals undergoing the GvH reaction^{5,6}. The suppression found in HvG reactions is also found in GvH reactions (Table 1, group E).

The specificity of the suppressive effect was studied by subcutaneous immunization of 'suppressed' mice with a combination of alloantigens comprising the antigen(s) used to induce the suppressor T cells as well as unrelated alloantigens. Reaction against the third party alloantigens was effectively suppressed, provided these antigens were presented in combination with the antigen(s) that had induced the suppressor T cells. This holds also for GvH reactions (Table 2). The phenomenon was observed for complete H-2 differences (groups A and C) as well as for H-2 subregion differences (groups B and D). Intravenous preimmunization with an alloantigen suppressed the DTH response to a challenge by completely different alloantigens, as long as the preimmunizing antigen was also present in the inducing inoculum (Table 2, groups A-C). The same result was obtained after suppression with non-H-2 alloantigens, subsequent immunization with a mixture of the same non-H-2 alloantigens plus unrelated 'bystander' H-2 antigens, and eventual testing for DTH reactivity against the H-2 antigens only (Table 3, groups A and B). These experiments clearly show that the 'bystander' antigens and the antigens used to induce the suppressor T cells do not have to be presented by the same cells.

Nonspecific suppressive effects have been demonstrated previously *in vivo*^{7,8} and *in vitro*⁹⁻¹³. Contact sensitizing agents^{14,15} as well as hapten-derivatized syngeneic^{16,17} and allogeneic¹⁸ cells readily induce suppressor T cell activity. The generation of this activity is dependent on complex interactions of three sets of suppressor T cells in a regulatory circuit under control of IgH and H-2 genes^{19,20}. Fresno *et al.*¹¹⁻¹³ have isolated an antigen-specific suppressor factor from continuously growing suppressor T cell clones. This factor was found to be a protein which, after interaction with antigen, breaks down into two peptides of 45,000 and 24,000 molecular weight. The former subunit suppresses both antigen-specific and other Lyt 1⁺ T cells.

Intravenous preimmunization with X-irradiated allogeneic cells also induces suppressor T cell activity. The type of the suppressor T cell activated depends on the dose of antigen and on treatment of the mice with cyclophosphamide²¹. Liew⁷, using high antigen doses and cyclophosphamide to induce suppressor T cell activity, has shown that suppression of DTH to H-2 subregion products can be induced if, and only if, the H-2 incompatibility includes the I-J subregion. The suppressor T cells are then antigen-specific, recognize the allo-I-J molecules and suppress also the DTH response to other H-2 subregion gene products if these are presented on the same cells as the allo-I-J determinants⁷. The suppressor T cells activated by our protocol (which differs from Liew's by using a lower dose of antigen and no cyclophosphamide) do not need to recognize allo-I-J determinants (Table 1) and will nonspecifically suppress the response to 'bystander' antigens, irrespective of whether the specific and 'bystander' antigens are physically associated.

Various mechanisms have been proposed to date for the beneficial effect of blood transfusion on kidney transplant survival^{22,23}. The data of Fresno *et al.*¹¹⁻¹³, combined with our finding of nonspecific suppression of the *in vivo* immune response to 'bystander' alloantigens, even if they are not

Table 1 Specificity of antigen recognition by suppressor T cells

System	Expt	Responding strain	Preimmunization i.v.	Immunization s.c.	% Specific increase foot thickness
HvG	A	BALB.B (H-2 ^d)	BALB.B (H-2 ^b)	BALB.B (H-2 ^b)	0
		BALB/c	BALB/c	BALB.B (H-2 ^b)	~25
		BALB/c	BALB.K (H-2 ^k)	BALB.B (H-2 ^b)	~25
	B	B10.AQR	B10.A (K)	B10.A (K)	~5
		B10.AQR	B10.AQR	B10.A (K)	~35
		B10.AQR	B10.T(6R)(I)	B10.A (K)	~35
	C	A.TH	A.TL (I)	A.TL (I)	0
		A.TH	A.TH	A.TL (I)	~25
		A.TH	A.SW (D)	A.TL (I)	~25
	D	B10.BR (D ^k)	B10.AKM (D ^q)	B10.AKM (D ^q)	~5
GvH	E	B10.BR	B10.BR	B10.AKM (D ^q)	~25
		B10.BR	B10.A(2R)(D ^b)	B10.AKM (D ^q)	~35
		B10.BR	B10.A(2R)(D ^b)	B10.A(2R)(D ^b)	~5
		B10.BR	B10.BR	B10.A(2R)(D ^b)	~25
		B10.BR	B10.AKM (D ^q)	B10.A(2R)(D ^b)	~25
		B10.BR	B10.AKM (D ^q)	B10.A(2R)(D ^b)	~25
		BALB/c (H-2 ^d)	BALB.B (H-2 ^b)	BALB.B (H-2 ^b)	~5
		BALB/c	BALB/c	BALB.B (H-2 ^b)	~25
		BALB/c	BALB.K (H-2 ^k)	BALB.B (H-2 ^b)	~25
					0 10 20 30 40

Suppressor T cells were induced by intravenous (i.v.) preimmunization of groups of six mice with 5×10^7 X-irradiated (2,000 rad) allogeneic spleen cells. HvG reactivity was stimulated 7 days later by subcutaneous (s.c.) injection of 1×10^7 allogeneic spleen cells, distributed over the inguinal area. Six days later the mice were tested for DTH reactivity by injection of 2×10^7 of the same allogeneic spleen cells into the dorsum of the right hind foot. Foot swelling was measured 24 h after challenge. Untreated control mice received only the challenge dose. The swelling in these control mice varied between 12 and 20%. DTH responses are expressed as % specific increase in foot thickness, and corrected for nonspecific swelling in the control mice. The histograms represent the arithmetic mean $\pm$ s.e. The H-2 haplotypes (group A), the H-2 subregion differences (groups B and C), and the haplotype of the H-2D locus (group D) are shown in parentheses. GvH reactions were elicited by i.v. injection of 1×10^7 spleen cells ('responder cells') into X-irradiated (700 rad) allogeneic recipient mice. Five days after reconstitution the total cell yield of spleen, inguinal, axillary and mesenteric lymph nodes of a recipient mouse was injected i.v. into an untreated secondary recipient syngeneic with the spleen cell donor. DTH reactivity was measured as in HvG reactions. The H-2 haplotype of the mice and cells used (group E) are shown in parentheses. The origin of the H-2 subregions (K, I-A, I-J, I-E, D): A.SW s s s s, A.TH s s s s, A.TL s k k k, BALB.B b b b b, BALB/c d d d d, BALB.K k k k k, B10.BR k k k k, B10.A k k k k, B10.AKM k k k k, B10.A(2R) q k k k, B10.AQR q k k k, B10.T(6R) q q q q.

Table 2 Specificity of the suppressive effect mediated by suppressor T cells

System	Expt	Responding strain	Preimmunization i.v.	Immunization s.c.	Challenge	% Specific increase foot thickness
HvG	A	BALB.B (H-2 ^b)	BALB/c (H-2 ^d)	BALB/c x K F ₁ (H-2 ^{d/k})	BALB/c x K F ₁ (H-2 ^{d/k})	~5
		BALB.B	BALB.B	BALB/c x K F ₁ (H-2 ^{d/k})	BALB/c x K F ₁ (H-2 ^{d/k})	~25
		BALB.B	BALB/c (H-2 ^d)	BALB/c x K F ₁ (H-2 ^{d/k})	BALB.K (H-2 ^k)	~5
		BALB.B	BALB.B (H-2 ^b)	BALB/c x K F ₁ (H-2 ^{d/k})	BALB.K (H-2 ^k)	~45
	B	B10.A	B10.AQR (K)	B10.T(6R) (K I)	B10.T(6R) (K I)	~5
		B10.A	B10.A	B10.T(6R) (K I)	B10.T(6R) (K I)	~25
		B10.AQR	B10.A (K)	B10.A x T(6R) F ₁ (K I)	B10.T(6R) (I)	~5
		B10.AQR	B10.AQR	B10.A x T(6R) F ₁ (K I)	B10.T(6R) (I)	~35
GvH	C	BALB.B (H-2 ^b)	BALB/c (H-2 ^d)	BALB/c x K F ₁ (H-2 ^{d/k})	BALB/c x K F ₁ (H-2 ^{d/k})	~5
		BALB.B	BALB.B	BALB/c x K F ₁ (H-2 ^{d/k})	BALB/c x K F ₁ (H-2 ^{d/k})	~25
		BALB.B	BALB/c (H-2 ^d)	BALB/c x K F ₁ (H-2 ^{d/k})	BALB.K (H-2 ^k)	~5
		BALB.B	BALB.B (H-2 ^b)	BALB/c x K F ₁ (H-2 ^{d/k})	BALB.K (H-2 ^k)	~25
	D	A.AL	A.TH (K I)	A.TH (K I)	A.TH (K I)	~5
		A.AL	A.AL	A.TH (K I)	A.TH (K I)	~35
		A.AL	ATL (K)	A.TH (K I)	A.TH (K I)	~5
						0 10 20 30 40 50

The experimental details are the same as in Table 1. The H-2 haplotypes (groups A and C), or the H-2 subregion differences (groups B and D) are shown in parentheses. BALB/c x K F₁ means F₁ hybrid of BALB/c x BALB.K, B10.A x T(6R) means F₁ hybrid of B10.A x B10.T(6R). The origin of the H-2 subregions of the A.AL strain is k k k k. The other strains are described in the legend to Table 1.

Table 3 Recognition of specific and 'bystander' antigens on separate cells

System	Expt	Responding strain	Preimmunization i.v.	Immunization s.c.	Challenge	% Specific increase foot thickness
A		B10.D2 (H-2 ^d)	BALB/c	BALB.B (H-2 ^b)	B10.ScSn (H-2 ^b)	
		B10.D2	B10.D2	BALB.B (H-2 ^b)	B10.ScSn (H-2 ^b)	
		B10.D2	BALB/c	BALB/c + B10.ScSn (H-2 ^b)	B10.ScSn (H-2 ^b)	
		B10.D2	B10.D2	BALB/c + B10.ScSn (H-2 ^b)	B10.ScSn (H-2 ^b)	
B		DBA/2 (H-2 ^d)	B10.D2	B10.G (H-2 ^q)	DBA/1 (H-2 ^q)	
		DBA/2	DBA/2	B10.G (H-2 ^q)	DBA/1 (H-2 ^q)	
		DBA/2	B10.D2	B10.D2 + DBA/1 (H-2 ^q)	DBA/1 (H-2 ^q)	
		DBA/2	DBA/2	B10.D2 + DBA/1 (H-2 ^q)	DBA/1 (H-2 ^q)	

0 10 20 30 40 50 60

Suppressor T cells against minor histocompatibility antigens were induced as described in the legend to Table 1, which also gives the details of the assays. The H-2 haplotype differences are shown in parentheses.

physically associated with the specific antigens, may well explain the blood transfusion effect. Sharing histocompatibility antigens by the transfused blood cells and the transplanted kidney and/or the passenger blood cells might reactivate antigen-specific suppressor T cells after transplantation, and that might suppress the anti-graft reactions in a nonspecific way.

We thank Drs S. Fazekas de StGroth, K. Fischer-Lindahl and M. H. Schreier for reading the manuscript, Dr Fazekas de

StGroth for help in the preparation of the manuscript, Mrs Lidia M. Hussaarts-Odijk and Mrs Beatrix D. Molendijk-Lok for technical assistance and Mrs Cary Meijerink-Clerkx and Mrs Jeanette van Dongen-Melman for typing the manuscript. This investigation was supported by grants from the Dutch Kidney Foundation, Amsterdam, and the Interuniversity Institute for Radiation Pathology and Radiation Protection (IRS), Leiden, The Netherlands.

Received 15 November; accepted 9 December 1982.

1. Van der Kwast, Th. H., Bianchi, A. T. J., Bril, H. & Benner, R. *Transplantation* **31**, 79-85 (1981).
2. Bril, H. & Benner, R. *Adv. exp. Med. Biol.* **149**, 577-583 (1982).
3. Bianchi, A. T. J., Hussaarts-Odijk, L. M. & Benner, R. *Transplant Proc.* (in the press).
4. Wolters, E. A. J. & Benner, R. *Nature* **279**, 642-643 (1979).
5. Wolters, E. A. J. & Benner, R. *Transplantation* **26**, 40-45 (1978).
6. Wolters, E. A. J. & Benner, R. *Transplantation* **27**, 39-42 (1979).
7. Liew, F. Y. *Eur. J. Immun.* **11**, 883-888 (1981).
8. Zembala, M., Asherson, G. L. & Colizzi, W. *Nature* **297**, 411-413 (1982).
9. Rich, S. S. & Rich, R. R. *J. exp. Med.* **140**, 1588-1602 (1974).
10. Rich, S. S. & Rich, R. R. *J. exp. Med.* **142**, 1391-1402 (1975).
11. Fresno, M., Nabel, G., McVay-Boudreau, L., Furthmayer, H. & Cantor, H. *J. exp. Med.* **153**, 1246-1259 (1981).
12. Fresno, M., McVay-Boudreau, L., Nabel, G. & Cantor, H. *J. exp. Med.* **153**, 1260-1274 (1981).

13. Fresno, M., McVay-Boudreau, L. & Cantor, H. *J. exp. Med.* **155**, 981-993 (1982).
14. Asherson, G. L. & Zembala, M. *Immunology* **42**, 1005-1013 (1980).
15. Claman, H. N., Miller, S. D., Conlon, P. J. & Moorhead, J. W. *Adv. Immun.* **30**, 121-157 (1980).
16. Miller, S. D., Sy, M.-S. & Claman, H. N. *J. Immun.* **121**, 265-273 (1978).
17. Bach, B. A., Sherman, L., Benacerraf, B. & Greene, M. I. *J. Immun.* **121**, 1460-1468 (1978).
18. Miller, S. D., Sy, M.-S. & Claman, H. N. *J. Immun.* **121**, 274-280 (1978).
19. Dietz, M. H. *et al. J. exp. Med.* **153**, 450-463 (1981).
20. Sy, M.-S., Nisonoff, A., Germain, R. N., Benacerraf, B. & Greene, M. I. *J. exp. Med.* **153**, 1415-1425 (1981).
21. Bianchi, A. T. J., Van der Kwast, Th. H. & Benner, R. *Immun. Today* **3**, 123-124 (1982).
22. Opelz, G., Sengar, D. P. S., Mickey, W. R. & Terasaki, P. I. *Transplant Proc.* **5**, 253-259 (1973).
23. Williams, K. A. & Morris, P. J. in *Organ Transplantation—Present State, Future Goals* (ed. Slavin, S.) (Elsevier, Amsterdam, 1980).

Cyclic AMP stabilizes a class of developmentally regulated *Dictyostelium discoideum* mRNAs

Giorgio Mangiarotti*†, Adriano Ceccarelli* & Harvey F. Lodish†

* Cattedra di Biologia Generale, Istituto di Igiene, University of Turin, Italy and

† Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The stability of mRNA is an important facet of the regulation of protein synthesis. In mammalian cells most mRNAs have long half-lives (5-15 hours) but a substantial fraction are much less stable^{1,2}. There are few examples where the stability of a particular mRNA or class of mRNAs is specifically affected by environmental or developmental stimuli. Certain hormones cause specific stabilization of mRNAs species³ and preferential mRNA stability is important in the accumulation of globin and myosin mRNAs during the terminal stages of erythropoiesis or myogenesis, respectively⁴⁻⁷. Disaggregation of *Dictyostelium discoideum* aggregates induces the specific destabilization of a large class of developmentally regulated mRNAs; thus, this system is an excellent one in which to determine how such controls are effected⁸⁻¹⁰. Here we show that addition of cyclic AMP to disaggregated cells specifically prevents the destabilization of these mRNAs.

During the aggregation stage of *Dictyostelium* differentiation, several thousand new mRNA species are transcribed and appear on polysomes—the 'aggregation-stage' mRNAs^{11,12}. Most mRNAs present in growing cells are also present in and synthesized by differentiating cells; these are termed 'common'

mRNAs. In the multicellular aggregates, both common and aggregation-stage mRNAs are equally stable, with half lives of about 4 h^{8,13}. If aggregates are disrupted and the cells shaken in suspension to prevent reaggregation, aggregation-stage mRNAs are rapidly lost from the cytoplasm, with a half life of 20 to 30 min, while the level of common mRNAs remains unaffected⁸⁻¹⁰. Addition of cyclic AMP to disaggregated cells prevents the loss of aggregation-stage mRNAs^{9,10}. Upon disaggregation, the rate of nuclear synthesis of aggregation-stage mRNAs is reduced by 5 to 10-fold, and addition of cyclic AMP to disaggregated cells restores transcription of aggregation-stage mRNAs to a normal or near normal rate¹⁴. However, the stimulatory effect of cyclic AMP on transcription of aggregation-stage mRNAs in disaggregated cells appears insufficient to explain the normal level of aggregation-stage mRNAs that is maintained in cells disaggregated in the presence of cyclic AMP.

In the first study that showed that addition of cyclic AMP prevents the destabilization of aggregation-stage mRNAs, cells were labelled with ³²P from 13 to 17 h of differentiation, after aggregation had occurred. At the end of the labelling period, unincorporated ³²P was removed by pressing the filters supporting the aggregated cells over a series of filter pads saturated with phosphate buffer, free of ³²P. The procedure was conducted for 1 h in order to deplete the internal pool of ³²P-labelled nucleotides. The cells were then disaggregated and shaken in suspension in the presence of 10 mM EDTA to prevent reaggregation. The cell suspension was divided in two equal portions, to one of which 100 μM cyclic AMP was added. At intervals during the chase, samples were withdrawn; and cytoplasmic polyadenylated RNA was purified. RNA was hybridized to a set of filter-bound cloned DNAs, under conditions⁸ where the ³²P radioactivity in RNA DNA hybrid is proportional to the

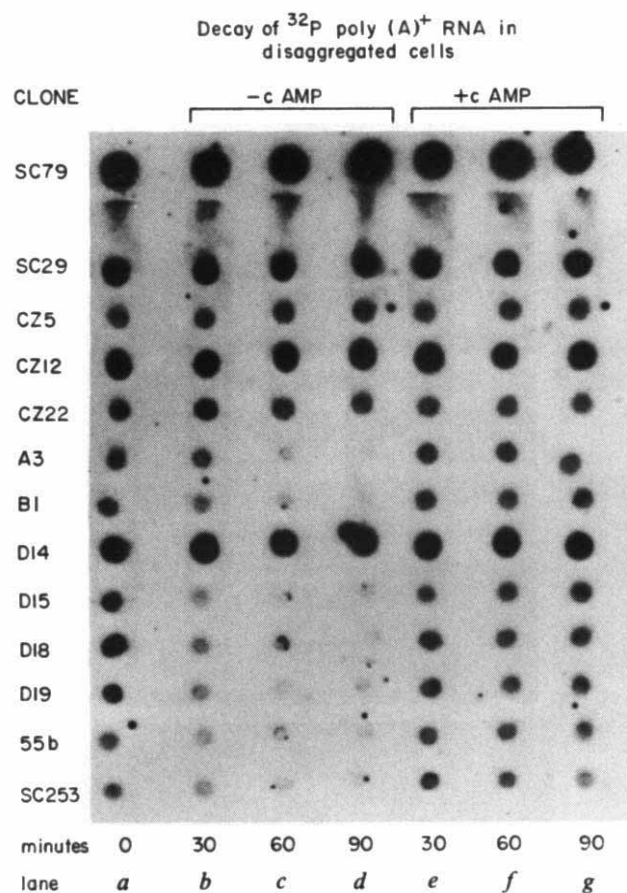


Fig. 1 Effect of cyclic AMP on the decay of ^{32}P label incorporated in polyadenylated RNA species following disruption of cell aggregates. Cells were grown, plated for development, labelled with ^{32}P from 13 to 17 h of differentiation and disaggregated⁸. After a zero-time sample was withdrawn from the cell suspension (lane a) the suspension was divided in two equal portions, to one of which 100 μM cyclic AMP was added. Samples were withdrawn at indicated times both from the suspension shaken in the absence (lanes b-d) and from the one shaken in the presence of cyclic AMP (lanes e-g). Cytoplasmic polyadenylated RNA was isolated and was hybridized to a set of filter-bound cloned DNAs as described in ref. 8 to measure the amount of label present in individual species of mRNA.

amount of ^{32}P mRNA species in the preparation. The amount of radioactivity in the common mRNAs—those complementary to clones SC79, SC29, CZ5, CZ12 and CZ22—was unaffected by disaggregation (Fig. 1, lanes a-d), and addition of cyclic AMP during the chase period had little effect on their levels (Fig. 1, lanes e-g). By contrast, disaggregation resulted in a rapid loss of radioactivity in all of the aggregation-stage mRNAs (those complementary to clones A3, B1, D15, D18, D19, 55b and SC253), with the sole exception of clone D14 (Fig. 1, lanes a-d). By scanning the radioautogram the half life of each of these mRNAs was determined to be about 30 min (data not shown). The loss of ^{32}P from these mRNAs during the chase period was prevented by addition of cyclic AMP (Fig. 1, lanes e-g), a result which indicates that cyclic AMP addition stabilizes the mRNAs in disaggregated cells.

Since addition of cyclic AMP to disaggregated cells restores normal synthesis of aggregation-stage mRNAs, some of the ^{32}P found in the aggregation-stage mRNAs during the presence of cyclic AMP could have been incorporated during the chase period. We feel that the incorporation of ^{32}P into RNA during the chase period is very low, if not nil, since ^{32}P radioactivity in rRNA, a relatively stable species with a half life of 10 to 12 h¹⁵, does not increase during the chase period. In fact, after a 90-min chase only 90% of the initial rRNA radioactivity

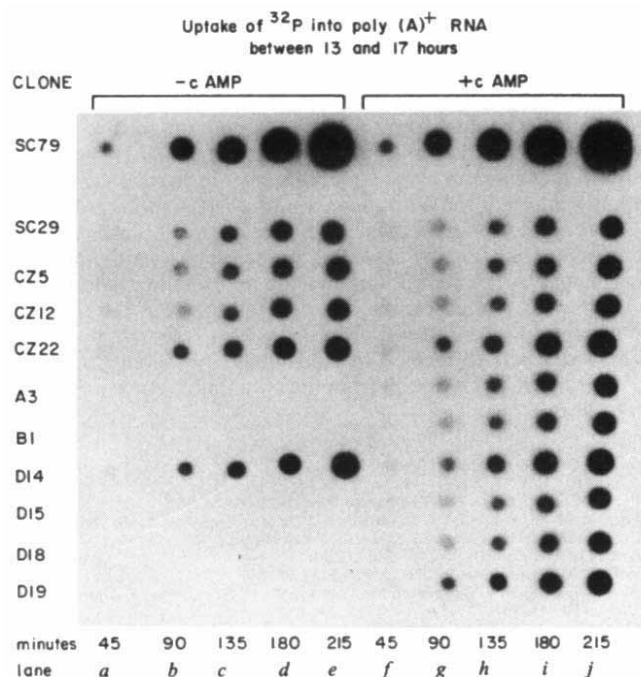


Fig. 2 Effect of cyclic AMP on the uptake of ^{32}P into cytoplasmic poly(A)⁺ RNA species in cells shaken in suspension following disaggregation. Cells at 15 h of development were disaggregated and shaken in suspension at a density of 5×10^5 cells per ml in the absence (lanes a-e) or in the presence of 100 μM cyclic AMP (lanes f-j). 10 MCI ^{32}P was added to each suspension at zero time. Samples of 30 ml were withdrawn at indicated times and cytoplasmic polyadenylated RNA was purified. The poly(A)⁺ RNA was hybridized to a set of filter-bound cloned DNAs as described in ref. 8 to measure the amount of label incorporated in individual species of mRNA.

remains. Transcription of common mRNAs and rRNA is unaffected by disaggregation¹⁴. We conclude that addition of cyclic AMP to disaggregated cells reversed the destabilization of aggregation-stage mRNAs.

To confirm this point, we performed a second type of experiment, approach to steady-state labelling. Cells were labelled with ^{32}P after they had been disaggregated; the uptake of the ^{32}P into specific polyadenylated mRNAs was followed while the cells were kept in suspension in the absence (Fig. 2, lanes a-e) or in the presence (Fig. 2, lanes f-j) of cyclic AMP. Whether or not cyclic AMP was added, ^{32}P was incorporated into all common mRNAs at a linear rate for at least 3.5 h. This means that the half life of these common mRNAs is at least 3.5 h (see ref. 8 for a discussion of these approach-to-steady-state labelling experiments). With the exception of clone D14, no detectable radioactivity was incorporated into any aggregation-stage mRNA by disaggregated cells incubated in the absence of cyclic AMP. Importantly, in the presence of cyclic AMP, ^{32}P was incorporated into all aggregation-stage mRNAs at a linear rate for at least 3.5 h. This means that, in cells disaggregated in the presence of cyclic AMP, the half-life of all aggregation stage mRNAs is at least 3.5 h. Were any of these mRNAs to decay more rapidly—say with a half life of 30 min—then the amount of ^{32}P in this species would plateau at 30–60 min; after this time incorporation of ^{32}P into this species would be balanced by degradation of ^{32}P -labelled mRNA.

We conclude that the addition of a high level of cyclic AMP to a suspension of disaggregated cells not only restores the synthesis of developmental mRNAs, but also prevents their destabilization. Whether cyclic AMP has a physiological role in controlling transcription and stability of mRNA in normal development is not known. Cyclic AMP is the extracellular chemotactic signal for cell-cell aggregation^{16,17}. This compound is also found at high levels in the multicellular aggregate¹⁸ and

influences the differentiation of prestalk and prespore cells¹⁹⁻²³. It is possible that the destabilization of aggregation-stage mRNAs on disaggregation of multicellular is due solely to a dilution of the concentration of extracellular cyclic AMP. Aggregating and post-aggregating cells contain surface cyclic AMP receptors. It is not known how the cyclic AMP signal is transduced across the plasma membrane¹⁷; in particular, we do not know whether the addition of cyclic AMP to disaggregated cells is reflected in an increase in the concentration of intracellular cyclic AMP. In any case, our studies reinforce the notion that changes in the stability of specific mRNAs in response to environmental stimuli can be an important facet in the regulation of protein biosynthesis.

This work was supported by an Italian CNR grant to G.M. and by grant PCM79-00839 from the National Science Foundation to H.F.L.

Received 11 October; accepted 17 December 1982.

- Puckett, L. & Darnell, J. E. *J. Cell Physiol.* **90**, 521-534 (1977).
- Harpold, M. M., Wilson, M. C. & Darnell, J. E. Jr *Molec. cell. Biol.* **1**, 188-198 (1981).
- Guyette, W. A., Matusik, R. J. & Rosen, J. M. *Cell* **17**, 1013-1023 (1979).
- Lodish, H. F. & Small, B. *Cell* **7**, 59-65 (1976).
- Volloch, V. & Houseman, D. *Cell* **23**, 509-514 (1981).
- Affara, N. A., Robert, B., Jacquet, M., Buckingham, M. E. & Gross, F. *J. molec. Biol.* **140**, 441-458 (1980).
- Buckingham, M. E., Caput, D., Cohen, A., Whalen, R. G. & Gross, F. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1466-1470 (1974).
- Mangiarotti, G., Lefebvre, P. & Lodish, H. F. *Dev. Biol.* **89**, 82-91 (1982).
- Landfear, S. M. & Lodish, H. F. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1044-1048 (1980).
- Chung, S., Landfear, S. M., Blumberg, D. D., Cohen, N. S. & Lodish, H. F. *Cell* **24**, 785-797 (1981).
- Blumberg, D. D. & Lodish, H. F. *Dev. Biol.* **81**, 74-80 (1980).
- Mangiarotti, G., Chung, S., Zuker, C. & Lodish, H. F. *Nucleic Acids Res.* **9**, 947-963 (1981).
- Margolske, J. P. & Lodish, H. F. *Dev. Biol.* **74**, 37-49 (1980).
- Landfear, S. M., Lefebvre, P., Chung, S. & Lodish, H. F. *Molec. cell. Biol.* **2**, 1417-1426 (1982).
- Mangiarotti, G., Altruda, F. & Lodish, H. F. *Molec. cell. Biol.* **1**, 35-42 (1981).
- Konijn, T. M., Van De Meene, J. G. C., Bonner, J. T. & Barkley, D. S. *Proc. natn. Acad. Sci. U.S.A.* **58**, 1152-1255 (1967).
- Devreotes, P. N. in *The Development of Dictyostelium discoideum* (ed. Loomis, W. F.) 117-168 (Academic, New York, 1982).
- Pahlic, M. & Rutheford, C. L. *J. biol. Chem.* **254**, 9703-9707 (1979).
- Bonner, J. T. *Proc. natn. Acad. Sci. U.S.A.* **65**, 110-113 (1970).
- Hamilton, I. D. & Chia, W. K. *J. gen. Microbiol.* **91**, 295-306 (1975).
- Town, C. D., Gross, J. D. & Kay, R. R. *Nature* **262**, 717-719 (1976).
- Town, C. & Stanford, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3228-3231 (1982).
- Kay, R. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3228-3231 (1982).

Maintenance of optical quality during crystalline lens growth

Russell D. Fernald & Stephen E. Wright

Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403, USA

Vertebrate lenses grow throughout life by the division of cells at the lens surface. The fibre cells thus produced are gradually covered by newer tissue, giving a layered structure. During growth, the lens must remain transparent and retain its refractile properties. The severity of these constraints is perhaps most evident in teleost fish which have a spherical lens that may increase in volume by a thousandfold during the first year of life¹. The dioptric power of the teleost fish eye is vested entirely in this spherical lens, as water, the cornea and the intraocular vitreous humour have almost identical refractive indices². Spherical lenses of uniform refractive index produce poor images because rays entering at different distances from the optic axis are focused at different distances from the lens. Teleost fish do not suffer from this imperfection and it has long been presumed that this is because there exists a refractive index gradient having a high value in the centre and decreasing continuously and symmetrically with radius in all directions^{3,4}. Here we demonstrate in the African cichlid fish, *Haplochromis burtoni*, that a refractive index gradient does exist, although its form is significantly different from that previously postulated⁵.

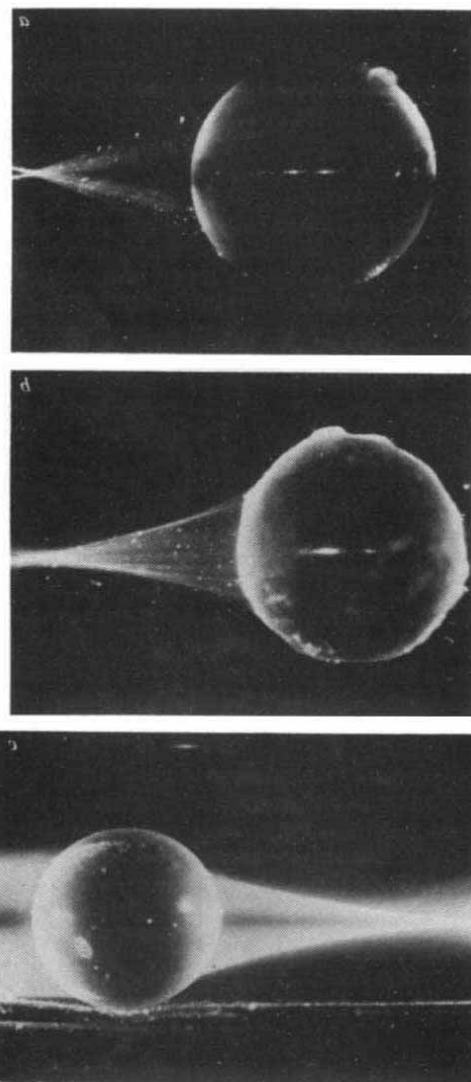


Fig. 1 *a*, Freshly excised *H. burtoni* lens ($r = 1.3$ mm) suspended by its ligament in fish Ringer's solution and illuminated by a laser beam. The finely focused cone of light is typical of illuminated fresh fish lenses. *b*, Same lens as in *a*, with the capsule and outer layers surgically removed, leaving the lens core which is illuminated by a laser beam. The edges of the focus cone of light are curved as is typical of lenses having a uniform refractive index. The core of the lens is firm, spherical and easily distinguished during surgery. After surgery, the lens core was placed briefly in a sonicator (Ultrasonics K42, 42 kHz) to remove tissue fragments on the exterior surface, as the surface debris produced significant photographic glare. Lens cores which were not sonicated had the same optical properties as those which were, and sonication of fresh whole lenses did not change their optical properties. *c*, For comparison, a glass sphere ($r = 2.47$ mm) of uniform refractive index ($n \sim 1.53$) showing the curved focus cone as in the lens core of *b*.

As the teleost, *H. burtoni*, is highly dependent on vision¹, it is startling that the diameter of the eye increases typically by a factor of 10 in the first year, with no adverse effects on visual function (R.D.F. and S.E.W., in preparation). We have analysed the refractive properties of the lens and suggest how the appropriate refractive power of the lens is maintained during growth.

We measured optical properties of the lens which are directly dependent on the refractive index, as direct measurement of refractive index is difficult in a lens known to have a refractive index gradient. Specimens ranging in size from 0.1 to 20 g were anaesthetized (MS222, 1:20,000; Sandoz) by immersion before enucleation. The lens was carefully removed from the excised

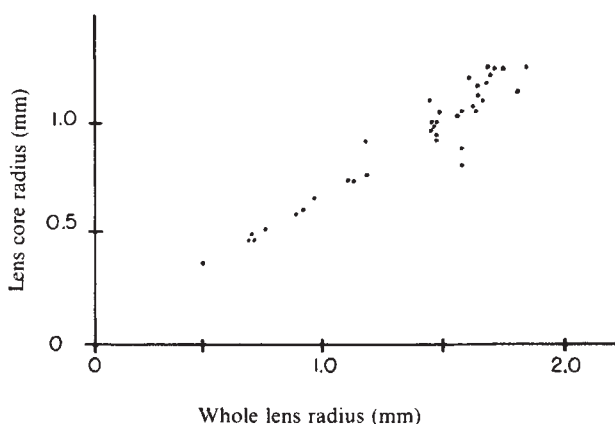


Fig. 2 The lens core radius as a function of whole lens radius; each point represents a separate lens. The high correlation coefficient ($r=0.96$) suggests highly regulated control of lens growth in *H. burtoni*.

eye and held by its suspensory ligament in a spectrophotometric cuvette containing fish Ringer's solution⁶ and 0.01% Ludox (Dupont H 540, refractive index in a solution of fish Ringer's = 1.334) to allow visualization of light beams in the cuvette. A laser beam (wavelength 494 nm; Coherent Radiation, Model 800) illuminated the lens and the cone of focus was photographed (see Fig. 1a). All of the freshly prepared lenses which we examined in this way, regardless of size, showed a finely focused cone of light terminating at a single point. After some time in the cuvette, however, a noticeable degradation of this focus occurred. Specifically, the outer cortex of the lens did not focus light as before, giving a broadly distributed focus. At this time, a lens core became visible and easily distinguishable within the lens. The lens capsule and outer layers of the lens were then surgically removed to isolate the core of the lens which was then placed in the chamber, held on a small platform (the suspensory ligament being no longer available), illuminated by the laser and photographed (Fig. 1b).

Lens cores focus light differently from whole lenses. First, no well defined focal point was produced by the isolated lens cores, but rather the emerging light rays assumed a shape characteristic of a spherically aberrant lens with a uniform refractive index. For comparison, Fig. 1c shows a laser beam focused by a glass bead. This bead, having a constant refractive index, also produces no single, well defined focal point but rather many foci, as rays nearer the optical axis (paraxial) are focused farther from the bead centre than are rays distant from that axis (marginal).

Second, the ratio of the paraxial focal length (f) to radius (r) of the lens core is significantly different from that of the intact lens. For the intact lens the ratio is $f/r=2.252$ (s.d. = 0.046, $N=12$, $\lambda=494$ nm), and for the core of the lens the ratio is $f/r=2.951$ (s.d. = 0.256, $N=10$). Matthiessen⁴ first measured this ratio in intact fish lenses and concluded that it was constant for all fish ($f/r=2.55$); it has since been called Matthiessen's ratio. Others^{7,8} have found values ranging from 2.10 to 2.49, depending on the species. Thus, Matthiessen's ratio depends on the species. Matthiessen's ratio is proportional to the refractive index as described below.

The difference between the types of tissue in the cortex and core of the lens is evident when surgically removing the cortex. The outer layers of the lens which are held in a spherical shape by the lens capsule are gelatinous and deteriorate quickly. Their

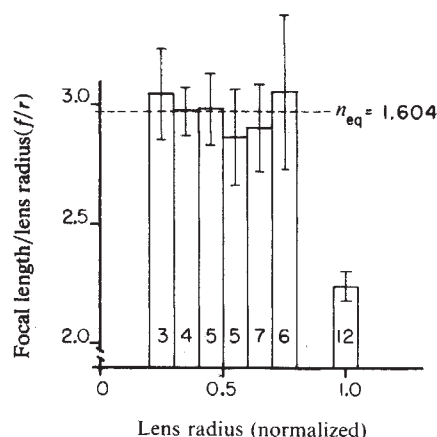


Fig. 3 The ratio of lens focal length to lens radius plotted as a function of normalized lens radius. The samples of 0.75r and smaller were obtained by surgically removing layers from the lens as described in the text. The width of each bar represents the range of sizes of the lens cores included in the sample (except of course, the whole lenses). Sample standard deviations are shown as bars on the histogram. If the refractive index of a lens is uniform, the ratio f/r is proportional to refractive index as follows: $n_{\text{lens}} = n_{\text{eq}}/[1 - (r/2f)]$. The fact that the values of f/r within the core are not different from one another suggests that there is no significant refractive index gradient within the lens core.

orderly, layered organization is evident only in freshly excised eyes or in well fixed preparations. In contrast, the lens core is a hard, nearly incompressible sphere, as reported previously for other teleosts^{2,9,10}. A lens core left in fish Ringer's for several days still exhibits the optical properties described above, whereas the lens cortex loses its optical integrity shortly after removal of the lens from the eye. We measured the relative sizes of this core and the cortex in 40 lenses and found that the radii of the core and cortex increase in direct proportion to the lens size (Fig. 2). The lens core radius is 0.674 (s.d. = 0.051, $N=40$) of the whole lens radius in all sizes of fish.

To assess the optical homogeneity of the core, we systematically removed additional layers from the lens core and repeated the above procedure, photographing the focused laser beams for each successively smaller lens core. This surgical 'peeling' of the lens was done with fine forceps under a dissecting microscope ($\times 12-25$). The clearly delineated lamina of the lens made it possible to obtain spherical lens cores with each 'peeling' by removing naturally occurring layers. By illuminating the lens with small laser beams, we measured the paraxial focal length of each lens core, which is related to the refractive index (n_{lens}) of the lens as¹¹:

$$n_{\text{lens}} = n_{\text{water}}/[1 - (r/2f)]$$

We found that f/r , and therefore the refractive index, remained unchanged as the core was made successively smaller (see Fig. 3).

Thus, for lenses of all sizes, the lens centre has a nearly uniform refractive index, and not a gradient as originally predicted^{12,13}. Instead, it appears that a steep refractive index gradient only exists between the outside of the lens core (0.67 radius, R) and the newly added cells at the outer edge of the lens (1.0 R). The newly added cells on the outside have a refractive index of about 1.38 (ref. 2), whereas the older, original cells nearer the lens centre have a refractive index close to 1.56 (ref. 4). This suggests that the maintenance of a nearly aplanatic lens during dramatic growth is achieved through the increasing thickness of the lens cortex which acts as a corrective coating on the spherically aberrant lens core.

Light rays passing through the lens cortex must be continuously curved if there is a refractive index gradient^{12,13}. To test this hypothesis directly, we must visualize a ray passing through the cortex, which is not possible in an intact lens as

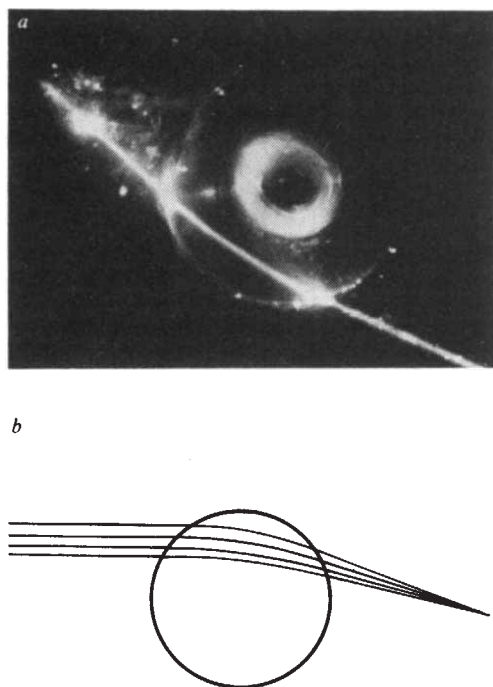


Fig. 4 *a*, A laser beam passing through a lens which has been cut in half to allow an undistorted view of the ray as it is refracted in the lens cortex. The continuous curve of the ray within the lens results from the radial gradient of refractive index. Attempts to photograph a similar ray in the core were unsuccessful due to glare from a cold cataract formed on thawing¹⁸. *b*, Tracing from several photographs similar to that in *a*, showing the paths of different rays. Rays nearer the lens core are less curved than those near the outer edge of the lens, suggesting that the refractive index gradient increases with radius. After enucleation an eye was quick-frozen onto a microtome stage and 10- μ m sections were taken at -18°C (Cryostat; IEC). On reaching the estimated lens centre, the frozen lens was removed from the eye and placed on a platform in the spectrophotometric cuvette. Illumination of the entire half lens revealed that the focal length was that predicted by the lens radius. The lens core remained opaque after thawing so that visualization of rays through the lens centre could not be done reliably (see also ref. 18).

the image of the ray would have to be viewed through the lens itself. To trace light rays passing through the cortex, we froze a whole eye and sectioned it on a microtome until we reached the centre of the lens (see Fig. 4 legend). The lens was then placed in a cuvette with the cut surface towards the camera, allowed to thaw, and rays passing through the lens cortex, just below the cut surface, were photographed as before. Figure 4*a* shows that a light ray passing through the cortex curves, confirming the existence of a steep refractive index gradient. By superimposing tracings from several photographs of different rays of light within the same lens (Fig. 4*b*), it is apparent that the rays of light are more strongly curved towards the edge of the lens, presumably because the gradient of refractive index increases with the radius.

On freezing, a cold cataract forms in fish lenses (at $\sim 0.44 R$; see Fig. 4*a*), so that laser beams passing through this region of the lens are scattered. Although no detectable gradient of refractive index exists in the core, the curvature of the laser beam as it passes through the outer layers of the core indicates that a slight refractive index gradient is present from $\sim 0.5 R$ to the cortex.

Based on these measurements, we suggest that the optical integrity of the fish lens is maintained during the growth of the fish eye as follows. As the lens grows, new cells are continually

added at its surface, extending around the entire circumference of the lens in layers of new tissue. This newly added lens tissue has a refractive index characteristic of live lens cells ($n \sim 1.38$). As new layers are added in a continuous process, cells in the older, more central layers slowly lose their water content and are compacted, joining the highly protein-concentrated core of the lens. With the loss of water, the refractive index increases to its maximum ($n \sim 1.56$), and no further increase in refractive index occurs. This accumulation of tissue in the centre of the lens results in a core of nearly uniform refractive index.

There is evidence that a similar process also occurs in other vertebrate lenses. In both the human¹⁴ and rat¹⁵ lens, the presence of a protein concentration gradient across the lens cortex suggests that there is a steep refractive index gradient also, as refractive index depends directly on protein concentration¹⁶. The lens core contains 60% protein whereas the cortex is only $\sim 20\%$ protein¹⁵. As protein and water make up almost 98% of the lens constituents¹⁵ this suggests that as the core forms, water is excluded from the lens cells, effectively increasing their protein concentration.

The increasing thickness of the cortical shell as the teleost lens grows, must have the proper refractive index gradient to correct the increasingly large spherically aberrant core. Regulation of both the lens size and the protein content of the cortex are important for the continued optical integrity of the lens and hence the eye during growth. The lens itself consists primarily of water and crystallin structural proteins. Crystallin lens proteins have been extracted and their properties (for example, molecular weight, charge and antigenic specificity¹⁷) compared both within species and within the lens of a single species. Generally, the individual proteins at the cortex surface are heteropolymers of high molecular weight, whereas those in the nucleus are monomers of low molecular weight. These differences are probably related to the corresponding configurational constraints on the structural proteins which allow tighter packing, higher density and thus higher refractive index towards the centre of the lens. Regulation of this transformation of protein structure within the lens during ontogeny is clearly critical for the optical integrity of the lens but the exact manner in which this control is exerted is as yet unknown.

This investigation was supported by a grant from the Whitehall Foundation. We thank Drs J. Presson, F. Munz and M. Westerfield for their helpful comments on the manuscript, H Howard for photographic work, N. Hirata and D. Wright for artwork.

Received 19 July 1982; accepted 4 January 1983.

1. Fernald, R. D. in *Advances in Neuroethology* (eds Ewert, J.-P., Capranica, R. R. & Ingle, D.) (Plenum, New York, in the press).
2. Hogben, L. & Landgrebe, F. *Proc. R. Soc. B* **128**, 317–342 (1938).
3. Maxwell, J. C. *Camb. Dubl. math. J.* **8**, 188–195 (1854).
4. Matthiessen, L. *Pflügers Arch. ges. Physiol.* **21**, 287–307 (1880); **27**, 510–523 (1882); **38**, 521–528 (1886).
5. Pumphrey, R. J. in *The Cell and the Organism* (ed. Ramsay, J. A.) 193–208 (Cambridge, London, 1961).
6. Palay, S. L., McGee-Russell, S. M., Gordon, S. Jr & Grillo, M. A. *J. Cell Biol.* **12**, 385–410 (1962).
7. Soemmering, D. W. *De Oculorum Hominis Animaliumque Secunde Horizontali Commentatio* (Vandenhoeck & Ruprecht, Göttingen, 1818).
8. Scroczynski, S. *J. comp. Physiol.* **121**, 135–144 (1977); *Zool. Jb. Physiol.* **79**, 204–212 (1975).
9. Brewster, D. *Phil. Trans. R. Soc.* 311–317 (1816).
10. Denton, E. J. *Bull. Mus. Hist. nat., Paris* **27**, 418–425 (1955); *Bull. Inst. Oceanogr., Monaco* **1071**, 1–10 (1956).
11. Sivak, J. G. *Vision Res.* **14**, 209–213 (1974).
12. Luneberg, R. K. *Mathematical Theory of Optics*, 208–213 (Brown University, Providence, 1944).
13. Fletcher, A., Murphy, T. & Young, A. *Proc. R. Soc. A* **223**, 216 (1954).
14. Fagerholm, P. P., Philipson, B. T. & Lindstrom, B. *Expl Eye Res.* **33**, 615–620 (1981).
15. Philipson, B. *Invest. Ophthalmol.* **8**, 258–270 (1969).
16. Doty, P. & Geiduschek, E. P. *The Proteins, Chemistry, Biological Activity and Methods* (eds Neurath, H. & Bailey, K.) 393–460 (Academic, New York, 1953).
17. Clayton, R. M. in *The Eye* Vol. 5 (eds Davson, H. & Graham, L. T. Jr) 399–494 (Academic, New York, 1974).
18. Bettelheim, F. A. & Lowenstein, M. A. *Expl Eye Res.* **28**, 651–663 (1979).

Phorbol esters increase the amount of Ca^{2+} , phospholipid-dependent protein kinase associated with plasma membrane

Andrew S. Kraft & Wayne B. Anderson*

Laboratory of Pathophysiology, National Cancer Institute, Building 10, Room B1B47, Bethesda, Maryland 20205, USA

Although the biochemical mechanism of action of phorbol ester tumour promoters is not fully understood, it is known that phorbol ester binding to the cell surface causes rapid changes in calcium flux^{1,2} and phospholipid metabolism³⁻⁷. A protein kinase activity has recently been described which is dependent on calcium and acidic phospholipids and is further enhanced by diacylglycerol⁸⁻¹⁰. Previously, we have observed that phorbol ester treatment of EL4 mouse thymoma cells causes a rapid decrease in cytosolic calcium, phospholipid-dependent protein kinase (Ca, PL-PK) activity, which is mediated through the specific phorbol ester cell-surface receptors identified on EL4 cells¹¹. We now show that treatment of parietal yolk sacs (PYS-2) cells with biologically active 12-*O*-tetradecanoyl phorbol-13-acetate (TPA) provokes a rapid decrease in cytosolic Ca, PL-PK activity that is accompanied by a significant increase in the amount of Ca, PL-PK activity associated with the plasma membrane fraction. These results suggest that the rapid and tight association of Ca, PL-PK activity with the plasma membrane may be an early event in mediating some of the effects of phorbol esters.

We have recently established that PYS cells contain a Ca, PL-PK activity similar to that described in other tissues¹². This PYS histone H₁ phosphotransferase activity is dependent on Ca^{2+} ($K_a \sim 32 \mu\text{M}$) and acidic phospholipids. The addition of diacylglycerol further stimulates Ca, PL-PK activity in the presence of Ca^{2+} and phospholipid. However, unlike other tissues examined, measurement of PYS Ca, PL-PK activity requires prior DEAE-cellulose chromatographic fractionation of the PYS supernatant. The chromatographic step serves to concentrate this kinase activity and to remove apparent endogenous inhibitors present in the crude supernatant preparation.

Figure 1a shows that passage of PYS 100,000g supernatant over a DEAE-cellulose column reveals Ca, PL-PK activity which elutes at 0.04 M NaCl. However, 30 min exposure of PYS cells to 0.1 μM TPA results in a pronounced decrease in the amount of Ca, PL-PK activity which can be detected in the cytosolic fraction (Fig. 1b). Direct addition of TPA to the reaction mixture in the presence of phosphatidylserine and diacylglycerol has no effect on cytosolic Ca, PL-PK activity (unpublished results). However, as recently described by Castagna *et al.*¹³, a direct activation of Ca, PL-PK *in vitro* can be demonstrated when TPA is added to a reaction mixture containing no diolefin. Thus, TPA is apparently able to replace the enzyme's requirement for unsaturated diacylglycerol. This effect of TPA on *in vitro* enzymatic activity may not be relevant to the TPA effects observed on Ca, PL-PK activity with treatment of intact cells. The calcium, phospholipid-dependent protein kinase contains hydrophobic regions, as indicated by the phospholipid requirement and purification properties¹⁴, and by the interaction of the enzyme with hydrophobic reagents such as antipsychotic drugs and local anaesthetics^{15,16}. Thus, any hydrophobic reagent (that is, phorbol ester) might interact with the enzyme to provoke either activation or inhibition.

To determine the time course of this phorbol ester-mediated loss in kinase activity, PYS cells were detached from dishes with a forceful stream of Dulbecco-Vogt modified Eagle's medium containing 10% fetal calf serum and incubated in

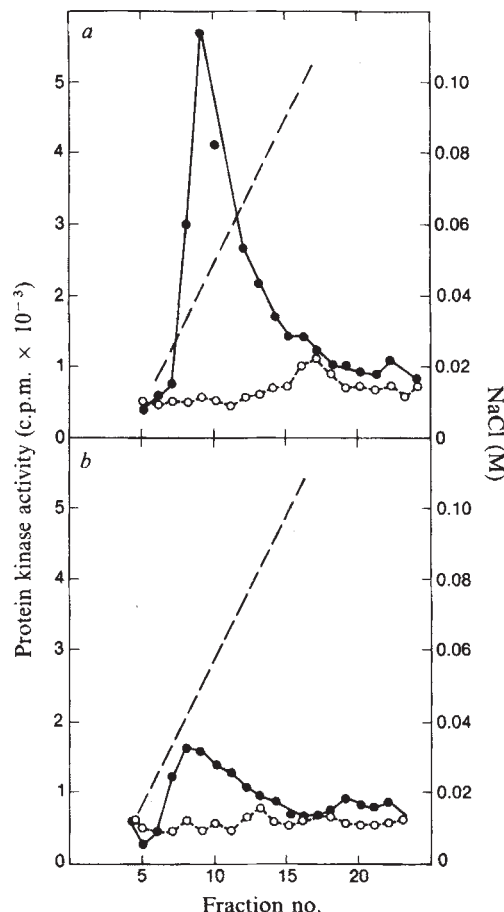


Fig. 1 Effect of TPA treatment of PYS cells on cytosolic Ca, PL-PK activity. PYS cells were grown on 150-mm plastic culture dishes in Dulbecco-Vogt modified Eagle's medium containing 10% fetal calf serum as described elsewhere¹². PYS cells were treated for 30 min at 37 °C with 0.1% ethanol (a) or with 0.1 μM TPA in 0.1% ethanol (b). After treatment the growth medium was removed and the cells (3.2×10^8 on eight dishes) were washed, collected, disrupted by Dounce homogenization and a 100,000g supernatant was prepared as previously described¹⁰. The 100,000g supernatants (11 mg protein each) prepared from control (a) and TPA-treated (b) cells were applied to identical DE52-cellulose columns ($0.9 \times 1.8 \text{ cm}$) equilibrated with buffer A (20 mM Tris-HCl pH 7.5, 0.5 mM EGTA, 2 mM EDTA, 2 mM phenylmethylsulphonyl fluoride). After sample addition the column was washed with 10 ml of buffer A, then eluted with a linear (0–0.15 M) NaCl gradient. Fractions of 1 ml were collected and 50- μl aliquots of each fraction were assayed for Ca, PL-PK activity as described elsewhere¹⁰ with 0.75 mM Ca^{2+} in the presence (●—●) or absence (○—○) of phospholipids (24 μg phosphatidylserine, 0.8 μg diolefin). Ca, PL-PK activity is determined by subtracting the activity measured in the absence of phospholipid from that measured in its presence. Protein kinase activity is expressed as c.p.m. ^{32}P incorporated per 3 min per 50- μl aliquot. ---, NaCl concentration.

suspension with TPA at 37 °C for various time periods up to 60 min. Within 5 min of TPA addition to intact PYS cells little Ca, PL-PK activity could be found in the cytosolic fraction, and activity remained very low throughout the entire period of TPA treatment. Although this effect of TPA seems to be very rapid, it must be noted that a further 30 min elapse while TPA-treated cells are washed, collected and disrupted before preparation of the 100,000g supernatant.

The loss of cytosolic Ca, PL-PK activity depends on the TPA concentration used to treat PYS cells (Fig. 3), treatment of the cells with 1 nM TPA for 30 min eliciting a ~20% decrease in cytosolic Ca, PL-PK activity, whereas 0.1 μM TPA causes a maximal decrease (~90%). Thus, the modulation of Ca, PL-PK activity occurs at TPA concentrations similar to those required

* Present address: Laboratory of Tumor Immunology and Biology, National Cancer Institute, Building 10, Room 8B02, Bethesda, Maryland 20205, USA.

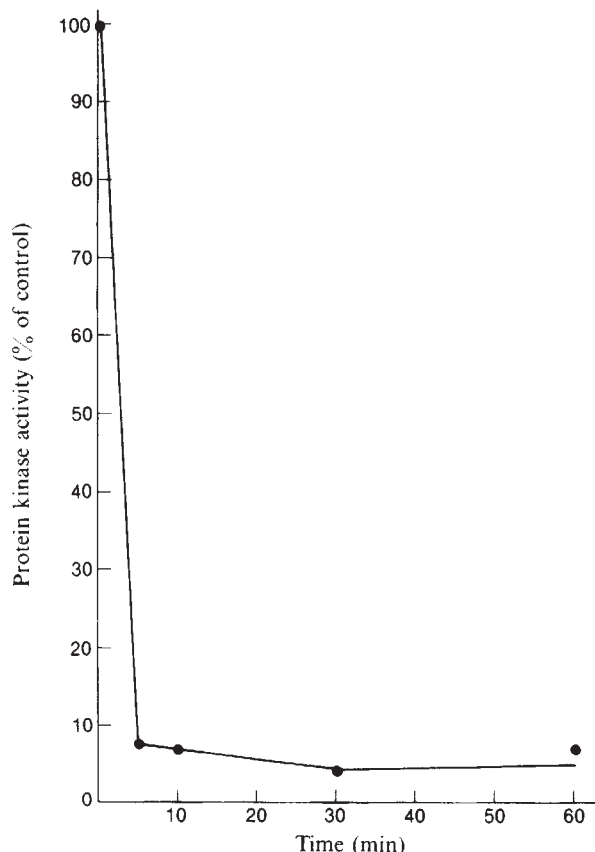


Fig. 2 Time course of the effect of TPA treatment of PYS cells on cytosolic Ca, PL-PK activity. PYS cells were washed from the dish with a forceful stream of growth medium and incubated in suspension at 37 °C in a shaking water bath. At time zero the cells in suspension were incubated with 0.1 μ M TPA. At the times indicated, aliquots of the treated cell suspension were removed, cells were washed and collected by centrifugation at 10,000g for 5 min at 4 °C and 100,000g supernatants prepared. The 100,000g supernatant (8 mg protein each) obtained at each time point was applied to a DEAE-cellulose column as described in Fig. 1 legend. After sample addition the column was washed with 10 ml of buffer A followed by 1.5 ml of 0.05 M NaCl in buffer A. Enzyme activity was then eluted batchwise with 3 ml of buffer A containing 0.05 M NaCl. This batch elution yields ~90% of the Ca, PL-PK activity observed with salt gradient elution. Aliquots (50 μ l) of the batch fractions were assayed for Ca, PL-PK activity as described in Fig. 1 legend. Results are expressed as the per cent of Ca, PL-PK activity found in the cytosol of cells treated with TPA compared with that found in cells treated with vehicle alone for the times indicated.

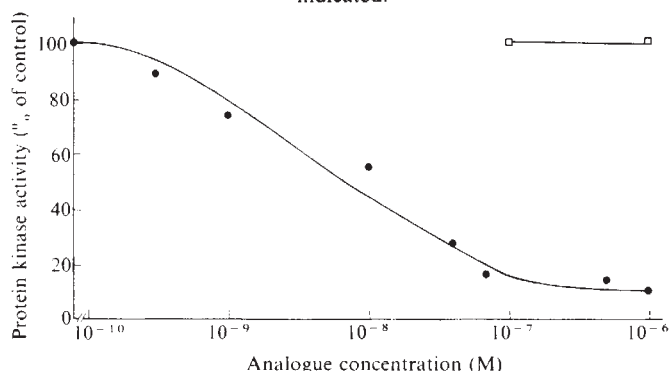


Fig. 3 Effect of treatment of PYS cells with increasing concentrations of phorbol esters on cytosolic Ca, PL-PK activity. PYS cells were treated with the indicated concentration of TPA (●) or 4- α -PDD (□) for 30 min at 37 °C. Supernatants (100,000g) were prepared as for Fig. 1 and the supernatants were fractionated by batchwise elution from DEAE-cellulose as described in Fig. 2 legend. Ca, PL-PK activity was determined and expressed as the per cent of untreated control as for Fig. 2.

to induce other biological effects of this tumour promoter. A decrease in cytosolic Ca, PL-PK activity is not observed when PYS cells are treated with the biologically inactive phorbol ester analogue, 4- α -phorbol 12,13-didecanoate (4- α -PDD) (Fig. 3).

The rapidity of the decrease in Ca, PL-PK activity following TPA addition suggested that, rather than a direct loss of activity, the enzyme might become associated with another subcellular fraction as a result of phorbol ester action. Ca, PL-PK activity has been demonstrated to be present in the particulate fraction of several mammalian tissues¹⁰. To measure the amount of Ca, PL-PK activity associated with the particulate fraction of PYS cells, the 100,000g pellet obtained from the crude homogenate was extracted with the nonionic detergent Nonidet P-40 (NP40, 1%) for 1 h at 4 °C. The detergent-treated preparation was centrifuged at 100,000g for 1 h and the NP40-solubilized material which remained in this 100,000g supernatant was fractionated by DEAE-cellulose chromatography. A small amount of Ca, PL-PK activity can be detected in the detergent-solubilized fraction obtained from untreated PYS cells (Fig. 4a). Interestingly, exposure of PYS cells to 0.1 μ M TPA for 30 min results in a dramatic increase in the amount of Ca, PL-PK activity found in the particulate fraction prepared from these TPA-treated cells (Fig. 4b). Quantitation of the amount of Ca, PL-PK activity lost from the cytosol with TPA treatment and the amount gained by the particulate fraction indicates that most of the lost activity can be accounted for by its association with the membrane fraction. Following TPA treatment of PYS cells for 30 min there is a loss of 3,040 units (1 unit = 1 pmol ³²P incorporated per 3 min per 10⁸ cells) of Ca, PL-PK activity from the cytosolic fraction and a gain of 2,640 units of Ca, PL-PK activity found associated with the particulate fraction (unpublished results).

A related study was carried out to determine if TPA treatment of PYS cells alters the amount of total (unfractionated cell homogenate) Ca, PL-PK activity. Crude homogenates were prepared from control and TPA-treated PYS cells and the total crude homogenate (combined cytosol plus particulate) was extracted with NP40. The resulting detergent-solubilized preparations were fractionated by DEAE-cellulose chromatography and the fractions assayed for Ca, PL-PK activity. Total activities of control and TPA-treated PYS cells agreed within $\pm$ 15%, again indicating that TPA treatment causes little change in total Ca, PL-PK activity. Rather, these results suggest that the rapid, and tight, association of Ca, PL-PK activity with the plasma membrane may be an early event in mediating some of the effects of phorbol esters (unpublished results).

To understand better the alteration in subcellular distribution of Ca, PL-PK activity following TPA treatment of PYS cells, crude subcellular fractions were prepared from TPA-treated PYS cells and assayed for this kinase activity. Each particulate fraction was extracted with NP40 and the detergent-solubilized preparations were then assayed for Ca, PL-PK activity. Negligible amounts of Ca, PL-PK were found in the cytosolic, nuclear and microsomal fractions, whereas the distribution of Ca, PL-PK activity found associated with the purified plasma membrane fraction (75% of activity) and the crude mitochondrial preparation (25% of activity) paralleled that of the plasma membrane enzyme marker adenylate cyclase (unpublished results).

The present results indicate that our initial observation that TPA treatment of intact cells provokes a marked reduction in cytosolic Ca, PL-PK activity is not limited to mouse thymocytes. Exposure of PYS cells to phorbol ester also results in a loss of cytoplasmic Ca, PL-PK activity which is time and TPA concentration dependent. We present evidence that TPA treatment can cause a substantial increase in the amount of Ca, PL-PK activity associated with the particulate fraction. Subcellular distribution studies indicate that this kinase activity becomes predominantly associated with the plasma membrane fraction. Attempts to dissociate this enzyme from the membrane fraction of TPA-treated PYS cells with high salt (0.3 M NaCl) buffer or with buffer containing metal chelators (1 mM EDTA and

EGTA) have been unsuccessful (unpublished results), suggesting that Ca, PL-PK has become an integral membrane component.

Conceivably, TPA treatment of cells might cause the translocation of Ca, PL-PK from the cytosol to the plasma membrane. Another possibility is that this enzyme normally exists within the cell in weak association with the plasma membrane and that TPA binding to the cell surface mediates a rapid change in membrane structure or composition to facilitate a tight interaction between Ca, PL-PK and the membrane. As TPA can interact directly with this protein kinase, it is conceivable that the enzyme might serve as, or be directly associated with, the TPA receptor. In this event, TPA treatment of intact cells may serve to stabilize the enzyme in association with the plasma membrane rather than mediate a rapid translocation from the cytosol to the membrane. However, we have been unable to demonstrate that TPA binding to isolated membranes might promote the rapid association of soluble Ca, PL-PK with the plasma membrane fraction.

Recent studies have shown that an initial event induced by growth stimulators¹⁷⁻²⁰ is enhanced phosphorylation of specific membrane proteins. The products of several different viral transforming genes have also been demonstrated to be protein kinases²¹⁻²⁴. The intracellular location of these transformation

protein kinases seems to be at the plasma membrane²⁵, and this membrane association is apparently necessary for *in vivo* tumour formation²⁶. Phorbol esters are also thought to act primarily at the cell surface²⁷. Thus, it is conceivable that the rapid TPA-induced association of Ca, PL-PK with the plasma membrane is an initial event in mediating eventual tumour formation. That growth promoters, viral transformation factors and the tumour promoter phorbol ester all might act through a mechanism of enhanced membrane phosphorylation perhaps suggests some degree of common substrate specificity.

Received 20 July; accepted 22 December 1982.

1. Gunther, G. J. *J. biol. Chem.* **256**, 12040-12045 (1981).
2. Verma, A. & Boutwell, R. *Biochem. biophys. Res. Commun.* **101**, 375-383 (1981).
3. Balman, A. & Hecker, E. *Biochim. biophys. Acta* **362**, 457-468 (1974).
4. Guy, G. & Murray, A. *Cancer Res.* **42**, 1980-1985 (1982).
5. Wertz, P. W. & Mueller, G. L. *Cancer Res.* **38**, 2900-2904 (1978).
6. Grove, R. I. & Schimmel, S. D. *Biochem. biophys. Res. Commun.* **102**, 158-164 (1981).
7. Ohuchi, K. & Levine, L. *J. biol. Chem.* **253**, 4783-4790 (1978).
8. Takai, Y. *et al. J. biol. Chem.* **254**, 3692-3695 (1979).
9. Kishimoto, A., Takai, Y., Mori, T., Kikkawa, U. & Nishizuka, Y. *J. biol. Chem.* **255**, 2273-2276 (1980).
10. Kuo, J. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7039-7043 (1980).
11. Kraft, A. S., Anderson, W. B., Cooper, H. L. & Sando, J. J. *J. biol. Chem.* **257**, 13193-13196 (1982).
12. Kraft, A. S. & Anderson, W. B. *J. biol. Chem.* (submitted).
13. Castagna, M. *et al. J. biol. Chem.* **257**, 7847-7851 (1982).
14. Wise, B. C., Raynor, R. L. & Kuo, J. F. *J. biol. Chem.* **257**, 8481-8488 (1982).
15. Mori, T., Takai, Y., Minakuchi, R., Yu, B. & Nishizuka, Y. *J. biol. Chem.* **255**, 8378-8380 (1980).
16. Schatzman, R. C., Wise, B. C. & Kuo, J. F. *Biochem. biophys. Res. Commun.* **98**, 669-676 (1981).
17. Ushiro, H. & Cohen, S. *J. biol. Chem.* **255**, 8363-8365 (1980).
18. Cohen, S., Carpenter, G. & King, L. Jr *J. biol. Chem.* **255**, 4834-4842 (1980).
19. Nilsen-Hamilton, M. & Hamilton, R. T. *Nature* **279**, 444-446 (1979).
20. Ek, B., Westermark, B., Wasteson, A. & Heldin, C.-H. *Nature* **295**, 419-420 (1982).
21. Erickson, R. L., Collett, M. S., Erickson, E. & Purchio, A. F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6260-6264 (1979).
22. Smith, A. E., Smith, R., Griffin, B. & Fried, M. *Cell* **18**, 915-924 (1979).
23. Witte, O. N., Dasgupta, A. & Baltimore, D. *Nature* **283**, 826-831 (1980).
24. Shih, T. Y., Papageorge, A. G., Stokes, P. E., Weeks, M. O. & Scolnick, E. M. *Nature* **287**, 686-691 (1980).
25. Hynes, R. O. *Cell* **21**, 601-602 (1980).
26. Krueger, J. G., Garber, E. A., Goldberg, A. R. & Hanafusa, H. *Cell* **28**, 889-896 (1982).
27. Cooper, R. A., Braunwald, A. D. & Kuo, A. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2865-2869 (1982).

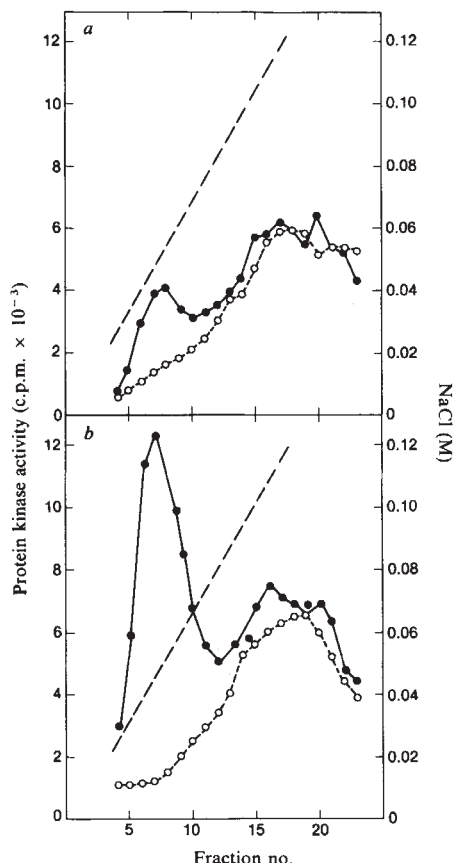


Fig. 4 Effect of TPA treatment of PYS cells on membrane-associated Ca, PL-PK activity. PYS cells were treated for 30 min at 37°C with 0.1% ethanol (a) or with 0.1 μ M TPA in 0.1% ethanol (b). Cells were collected and homogenized and a 100,000g pellet was prepared as for Fig. 1. The resulting pellets were resuspended in 5 ml of buffer A containing 1% NP40 and Dounce homogenized (30 strokes). This detergent-treated suspension was slowly rotated at 4°C for 60 min and then centrifuged at 100,000g for 60 min. The detergent-solubilized supernatant fraction was applied to a DEAE-cellulose column and eluted with a 0-0.15 M salt gradient as described in Fig. 1 legend. Fractions of 1 ml were collected and assayed for Ca, PL-PK activity with 0.75 mM Ca^{2+} in the presence (●—●) or absence (○—○) of phospholipids. Protein kinase activity is expressed as for Fig. 1. ---, NaCl concentration.

Structure of the serine chemoreceptor in *Escherichia coli*

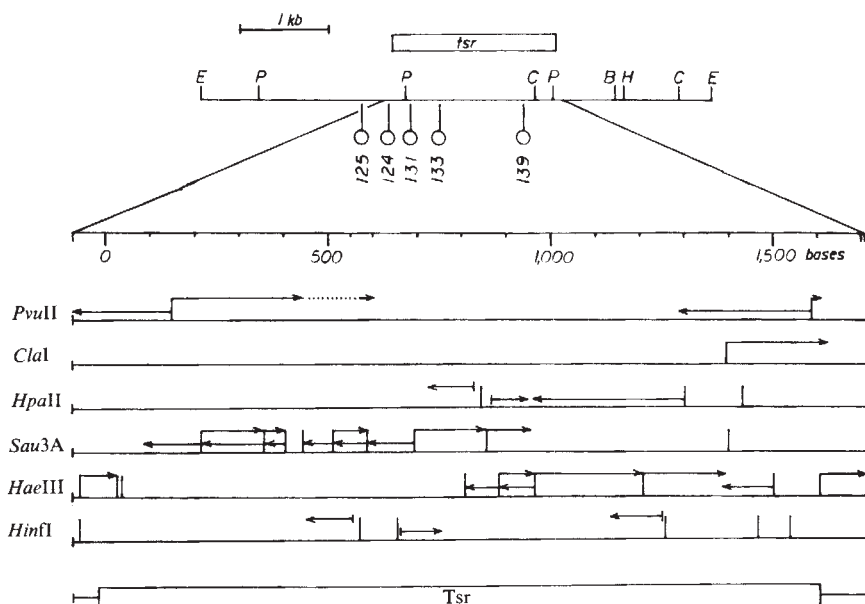
Alan Boyd, Katherine Kendall & Melvin I. Simon

Department of Biology, B-022, University of California, San Diego, La Jolla, California 92093, USA

Many biological processes depend on the function of proteins that detect changes in a cell's environment and transmit the information to the cytoplasm, for example, peptide hormone receptors¹. In *Escherichia coli* this class of proteins is exemplified by the sensory transducers (also called signalling proteins or methyl-accepting chemotaxis proteins) which have a central role in mediating chemotactic behaviour^{2,3}. The sensory transducers are the products of four genes: *tsr*, *tar*, *tap* and *trg*. Each transducer detects changes in the environmental concentration of one or a very few attractants: Tsr, serine; Tar, aspartate and maltose; Tap, unknown; and Trg, ribose and galactose. Tsr and Tar act directly as chemoreceptors for the amino acid attractants^{4,5} and signal changes in their degree of occupancy to the flagellar apparatus. Detection of these changes in occupancy is made possible as the transducers are methylated at multiple glutamate residues⁶⁻⁹ such that their level of methylation reflects the most recent chemoeffector concentration. Biochemical^{10,11} and genetic^{12,13} information concerning the serine transducer protein has been accumulating rapidly but little is known about the structure of the molecule. We present here the nucleotide sequence of the *tsr* gene of *E. coli*; the amino acid sequence derived from it suggests that the Tsr transducer protein has a relatively simple transmembrane structure that may place limits on the mechanisms available for the transmission of sensory information into the cell.

We have described previously how the kanamycin resistance transposon Tn5 was used to map the *tsr* gene carried by the

Fig. 1 Strategy for sequencing the *tsr* gene. Upper: the 5.5-kilobase (kb) *EcoRI* insert fragment of pAB100 showing the positions of several insertions of Tn5 used to define the *tsr* region¹². The pAB100 derivative carrying Tn5-124 complemented a *tsr* defect, but did so more weakly than other plasmids carrying a Tn5 insertion well outside the *tsr* region (for example, Tn5-125). Restriction mapping localized Tn5-124 to the *tsr* side of the *Bst*N1 site at nucleotide -40 (Fig. 2). Furthermore, one of the novel *Bst*N1 fragments created by this insertion was found to be detectably larger than the target fragment. As Tn5 has symmetrical *Bst*N1 sites located 49 base pairs (bp) from each end²⁸ it seems that Tn5-124 lies considerably less than 49 bp to the right of the *Bst*N1 site at -40, and that the insertion has disrupted sequences necessary for normal levels of *tsr* expression, leading to reduced complementation. E, *EcoRI*; P, *PvuII*; C, *Clal*; B, *BglII*; H, *HindIII*. Lower: the region of the pAB100 insert fragment that was sequenced. A map of restriction sites was determined extending leftwards from the *BglII* site by the procedure of Smith and Birnstiel²⁹. Sequencing runs using the methods of Maxam and Gilbert³⁰ are indicated by arrows ending in vertical slashes (from *HpaII* and *HinfI* sites). These were performed with DNA fragments labelled at their 3' ends by *E. coli* DNA polymerase I (Klenow fragment). The rest of the sequence was determined by the dideoxy procedure³¹, using M13mp7 recombinant phage plus-strand DNA as template³². Initially, two libraries of clones were constructed by digesting gel-purified 5.5-kb *EcoRI* fragment with either *Sau3A* or *HaeIII* and ligating the fragments into M13mp7 replicative form (RF) DNA which had been digested with *Bam*HI or *Hinc*II, respectively. A combination of the following strategies was then used to identify clones for sequencing: (1) viral plus-strand DNA was blotted from agarose gels to nitrocellulose and probed successively with several restriction fragments; (2) the C-test procedure³¹ for identifying M13 clones bearing fragments having complementary sequences was used to screen for complementary clones within a library, and to find overlapping clones in the other library. Then, *PvuII* fragments of pAB100 were ligated into *Hinc*II-cleaved M13mp7 RF; *Clal* fragments of pAB100 were ligated into *AccI*-cleaved M13mp7 RF and a gel-purified *HpaII* fragment was ligated into *AccI* cleaved M13mp7 RF. Both of the *Clal* sites in the pAB100 insert were cleaved only when DNA was prepared from a bacterial strain bearing a *dam* mutation. Arrows denote the direction and extent of sequencing. The dotted arrow in a *PvuII* sequencing run indicates that a distal portion of that particular gel was used to read across a *Sau3A* site.



recombinant plasmid pAB100 (ref. 12). To determine the amino acid sequence of the Tsr protein we determined the nucleotide sequence of the genetically defined *tsr* region (see Fig. 1). This sequence of 1,788 nucleotides (Fig. 2) includes a long open reading frame which spans the positions of three well characterized *tsr* :: Tn5 insertions. A leftward limit for the start of the *tsr* gene is defined by the Tn5 insertion numbered 124 in Fig. 1 (Tn5-124). Within this leftward limit the assignment of the start of the *tsr* gene is relatively unambiguous. There are no in-phase initiation codons lying between the ATG assigned as the *tsr* start in Fig. 2 and the *Bst*N1 site at nucleotide -40. Downstream from the *tsr* start there are no other ATG codons until that encoding Met 75. Upstream from the proposed initiation codon are two overlapping sequences that could act as ribosome-binding sites¹⁴ (-12 to -4 and -15 to -8); we have not attempted to identify potential transcriptional promoters. The open reading frame is terminated by a single TGA codon. There is no obvious rho-independent terminator¹⁵ downstream from *tsr*; a stable stem-loop could be formed by nucleotides 1,667-1,696 if this region were transcribed, but this is not followed by a run of T residues.

The open reading frame identified as the structural gene for Tsr should encode a primary polypeptide product consisting of 536 amino acids (with no cysteine residues), and having a calculated molecular weight (MW) of 57,483, which agrees with the electrophoretic mobility of Tsr⁶. There are more acidic than basic residues, which agrees with the observed isoelectric point of Tsr¹⁰. Although the overall amino acid composition of Tsr is not markedly rich in hydrophobic residues (53% charged plus uncharged polar) there are two notable stretches of predominantly hydrophobic amino acids: Ile 7 to Leu 33 and Asn 185 to Ile 214. Both of these have high indices of hydrophobicity (2.2 and 2.5, respectively) calculated according to Segrest and Feldman¹⁶, suggesting that they could span a lipid bilayer. The former sequence, being located at the amino-terminus and preceded by a cluster of basic residues, closely resembles the signal sequences of several proteins that are

exported to the periplasm and outer membrane of *E. coli*¹⁷. Such signal sequences are thought to initiate the extrusion of a growing polypeptide chain across the cytoplasmic membrane, after which they are proteolytically removed. Although there are several potential processing sites in the putative signal sequence of Tsr, it is unknown whether the protein is processed during assembly.

The sequences of several bacterial membrane proteins are now known: bacteriorhodopsin¹⁸, lactose permease¹⁹, the P, Q and M components of histidine permease²⁰, and NADH dehydrogenase²¹. In contrast to Tsr, these proteins lack signal sequences and are generally very hydrophobic. These membrane proteins all act as permeases and presumably have a complex transmembrane structure like bacteriorhodopsin, which traverses the membrane seven times²². The Tsr sequence is more akin to the sequences of several transmembrane proteins found in eukaryotes, for example, the G protein of vesicular stomatitis virus and the haemagglutinin of influenza virus. Insertion of these proteins into the membrane is initiated by a signal sequence and continues until a second hydrophobic amino acid sequence enters the membrane (reviewed in ref. 23). The second hydrophobic sequence is thought to act as a stop-transfer signal, and remains within the membrane, anchoring the protein in position. The contrast between the membrane proteins of eukaryotes and prokaryotes with respect to their modes of assembly may merely reflect an underlying bias in the proteins chosen for study—'recognition proteins' in eukaryotes and 'transport proteins' in prokaryotes. The primary structure of Tsr indicates that it may represent a prokaryotic 'recognition protein'. The sequence Asn 185 to Ile 214 might function as a stop-transfer signal. There are no other significant stretches of hydrophobic amino acids between the putative signal sequence and the sequence Asn 185-Ile 214 (the tract of uncharged residues Gln 48-Thr 63 has a hydrophobicity index of only 0.97). According to this interpretation of the sequence, the amino-terminal domain of Tsr is exported to the periplasmic face of the membrane while the carboxy-terminal

-60. -50. -40. -30. -20. -10. -1.
TTATAAGTTTTTCCTTCCAGGCCGAAATCTTCACGCTCCACGGAAGACAAC

1
met leu lys arg ile lys ile val thr ser leu leu leu val leu ala val phe gly leu leu gln leu thr ser gly gly leu phe phe
ATG TTA AAA CGT ATC AAA ATT GTG ACC AGC TTA CTG CTG GTT TTG GCC GTT TTT GGC CTT TTA CAA CTG ACA TCA GGC GGT CTG TTC TTT

31
asn ala leu lys asn asp lys glu asn phe thr val leu gln thr ile arg gln gln gln ser thr leu asn gly ser trp val ala leu
AAT GCC TTA AAG AAT GAC AAA GAA AAT TTC ACT GTT TTA CAA ACC ATT CGC CAG CAG CAA TCC ACG CTG AAT GGC AGC TGG GTC GCG TTG

61
leu gln thr arg asn thr leu asn arg ala gly ile arg tyr met met asp gln asn asn ile gly ser gly ser thr val ala glu leu
TTG CAG ACC CGT AAC ACC CTC AAC CGC GCG GGT ATC CGC TAC ATG ATG GAT CAG AAT AAT ATT GGT AGC GGT TCA ACC GTT GCT GAG CTG

91
met glu ser ala ser ile ser leu lys gln ala glu lys asn trp ala asp tyr glu ala leu pro arg asp pro arg gln ser thr ala
ATG GAG AGT GCC AGT ATT TCG CTG AAA CAG GCG GAA AAA AAC TGG GCG GAT TAC GAA GCG TTG CCG CGT GAC CCG CGT CAG AGC ACC GCG

121
ala ala ala glu ile lys arg asn tyr asp ile tyr his asn ala leu ala glu leu ile gln leu leu gly ala gly lys ser thr ser
GCA GCG GCA GAG ATC AAA CGT AAT TAC GAT ATT TAT CAC AAT GCG CTG GCG GAG CTG ATC CAA CTG TTA GGT GCA GGC AAA TCA ACG AGT

151
ser leu ile ser arg pro arg asp ile arg asn gly phe glu lys gln tyr val ala tyr met glu gln asn asp arg leu his asp ile
TCT TTG ATC AGC CGA CCC AGG GAT ATC AGG AAC GGT TTC GAG AAG CAG TAT GTG GCT TAC ATG GAG CAA AAC GAT CCG CTC CAT GAT ATC

181
ala val ser asp asn asn ala ser tyr ser gln ala met trp ile leu val gly val met ile val val leu ala val ile phe ala val
GCC GTC AGC GAT AAC AAT GCC TCC TAC AGC CAG GCG ATG TGG ATT CTG GTG GCG GTG ATG ATC GTC GTA CTG GCG GTC ATC TTC GCC GTC

211
trp phe gly ile lys ala ser leu val ala pro met asn arg leu ile asp ser ile arg his ile ala gly gly asp leu val lys pro
TGG TTC GGT ATT AAA GCG TCG CTG GTA GCG CCA ATG AAT CCG CTG ATT GAC AGC ATT CGT CAT ATT GCA GCG GCG GAT CTG GTG AAA CCG

241
ile glu val asp gly ser asn glu met gly gln leu ala glu ser leu arg his met gln gly glu leu met arg thr val gly asp val
ATT GAG CTG GAT GGC TCT AAT GAG ATG GGG CAA CTG GCA GAG AGT TTG GCG CAT ATG CAG GGA GAG CTG ATG CGT ACC GTC GGT GAT GTG

271
arg asn gly ala asn ala ile tyr ser gly ala ser glu ile ala thr gly asn asn asp leu ser ser arg thr gly gln gln ala ala
CGC AAC GGG GCG AAT GCC ATC TAT AGC GGT GCC AGC GAA ATC GCC ACC GGC AAT AAC GAT CTC TCT TCG CGC ACC GAG CAA CAG GCC GCT

301
ser leu glu glu thr ala ala ser met glu gln leu thr ala thr val lys gln asn ala glu asn ala arg gln ala ser his leu ala
TCG CTG GAA GAG ACG GCA GCG AGC ATG GAG CAA CTG ACC GCA ACG GTG AAG CAG AAC GCG GAA AAT GCG CGC CAG GCC AGC CAT CTG GCG

331
leu ser ala ser glu thr ala gln arg gly gly lys val val asp asn val val gln thr met arg asp ile ser thr ser ser gln lys
TTA AGT GCT TCT GAA ACG GCG CAA CCG GCG GGT AAA GTG GTA GAT AAC GTG GTG CAG ACT ATG CCG GAT ATC TCC ACC AGT TCG CAG AAA

361
ile ala asp ile ile ser val ile asp gly ile ala phe gln thr asn ile leu ala leu asn ala ala val glu ala ala arg ala gly
ATC GCC GAT ATT ATC AGC GTA ATT GAC GCG ATT GCC TTC CAG ACC AAT ATT CTG GCT TTG AAC GCG GCG GTT GAG GCT GCG CGT GCG GGT

391
glu gln gly arg gly phe ala val val ala gly glu val arg asn leu ala gln arg ser ala gln ala ala arg glu ile lys ser leu
GAG CAA GGG GCG GGT TTT GCG GTG GTC GCG GGA GAA GTG CGT AAT CTG GCC CAG GCG AGC GCC CAG GCG GCT CGT GAA ATT AAA AGC CTG

421
ile glu asp ser val gly lys val asp val gly ser thr leu val glu ser ala gly glu thr met ala glu ile val ser ala val thr
ATT GAA GAC TCG GTG GGG AAA GTG GAT GTT GGC TCT ACG CTG GTC GAA AGC GCC GGG GAA ACA ATG GCG GAG ATT GTC AGC GCC GTG ACC

451
arg val thr asp ile met gly glu ile ala ser ala ser asp glu gln ser arg gly ile asp gln val gly leu ala val ala glu met
CGC GTG ACG GAC ATT ATG GGC GAA ATT GCT TCT GCT TCT GAT GAG CAG AGC CGT GGT ATC GAT CAG GTT GGC TTA GCG GTT GCT GAG ATG

481
asp arg val thr gln gln asn ala ala leu val glu glu ser ala ala ala ala ala leu glu glu gln ala ser arg leu thr glu
GAC CCG GTA ACT CAA CAG AAC GCC GCG CTG GTG GAA GAG TCT GCC GCT GCC GCC GCG CTG GAA GAG CAG GCC AGT CCG CTG ACC GAA

511
ala val ala val phe arg ile gln gln gln gln arg glu thr ser ala val val lys thr val thr pro ala ala pro TER 1620.
GCA GTG GCA GTG TTC CCG ATT CAG CAA CAG CAG CGT GAA ACA TCG GCT GTG GTA AAA ACC GTG ACG CCA OCT CCG CCG TGA AAATGGCCCTG

1630. 1640. 1650. 1660. 1670. 1680. 1690. 1700. 1710. 1720.
CGAGATAGCGAGGAGAACTGGGAACATTTTAATCGCCATGAAATGCCGATAAGCAAAATGTTATCGGGCATAAGGAGATTAATCTTTACGTGGGTCGTTGATC

Fig. 2 Nucleotide sequence of the *tsr* gene. A sequence of 1,788 nucleotides is shown, including the entire *tsr* structural gene. The strand shown is the *tsr* anticoding strand. The derived amino acid sequence for the Tsr protein is also presented. Numbers within the *tsr* structural gene refer to amino acid positions, with dots at every tenth residue. In the flanking sequences the numbering refers to nucleotides, position 1 being the first base of the initiation codon, and position -1 being the first base upstream. 75% of the sequence was determined on both strands of the DNA, and all single-stranded regions, were sequenced in at least two separate experiments. Sequence data were manipulated and analysed using the computer programs of Staden³³.

domain from Lys 215 onwards is constrained to fold inside the cell. Evidence consistent with this transmembrane conformation of Tsr comes from the finding that tryptic digestion of Tsr in intact bacterial spheroplasts is limited to the removal of ~200 amino acids of the molecule (A.B., unpublished data) indicating that the bulk of the protein is inaccessible to external proteases.

The functionally analogous serine and aspartate-maltose transducers are encoded by genes exhibiting strong nucleotide sequence homology in their 3' segments¹². This finding led to the proposal that the carboxy-terminal domains of the proteins contain structurally homologous sites of methylation involved in sensory adaptation. During sensory adaptation, Tsr is subject to methylation and demethylation at the carboxyl groups of five or six glutamic acid residues. These modifications are performed by the CheR methyltransferase and the CheB methyl-esterase, both of which are located in the cytoplasm (reviewed in ref. 3). Kehry and Dahlquist¹⁰ have identified and characterized two methyl-labelled peptides in tryptic digests of Tsr. The chromatographic properties of these peptides indicate that they are large and/or relatively nonpolar. One peptide, R1 (nomenclature of ref. 24), is an arginine-containing peptide with no methionines that carries two sites of methylation; the other peptide, K1, is a lysine-containing peptide with a single methionine at position 15, which carries three or four sites of

methylation. The tryptic peptides predicted from the sequence of Tsr include only two peptides that meet all of these criteria (see Fig. 3a); the two peptides exhibit some homology that does not, however, extend beyond the sequences shown. Both lie in the carboxy-terminal portion of the molecule, consistent with the proposal that the membrane-spanning peptide delineates extra- and intracytoplasmic domains. The positions of four sites of methylation in the K1 and R1 peptides have been determined²⁴; these are shown in Fig. 3, together with a possible fifth site which we propose on the basis of sequence homology. The methylation sites identified all occur at the second position in a pair of glutamate residues (after CheB deamidation). Two of the methylation sites correspond to residues encoded as glutamine in the *tsr* gene. This provides strong support for the proposal that an irreversible covalent modification performed by the CheB methyl-esterase^{25,26}, which in some way increases the availability of methyl-accepting sites, is the deamidation of glutamine residues. The role of this second modification in sensory transduction is not clear, but the finding that it involves the unmasking of glutamate residues destined to be methylated during sensory adaptation suggests that it may be a mechanism whereby newly synthesized transducer proteins can be introduced into the cell membrane in an inactive or attenuated form.

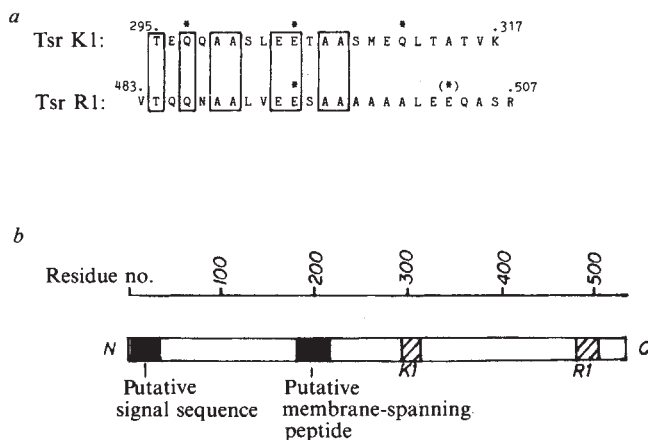


Fig. 3 *a*, Amino acid sequences of the methylated regions of Tsr. The sequences of two tryptic peptides derived from the *tsr* DNA sequence are shown, aligned to indicate internal homologies. The nomenclature K1 and R1 is that of ref. 24. * denotes sites of carboxymethylation as determined by peptide sequencing²⁴; (*) denotes a further site in R1 proposed here on the basis of sequence homology. *b*, Schematic diagram of Tsr primary structure. Solid regions are stretches of hydrophobic sequence; cross-hatched regions indicate the positions of the methylated peptides.

Figure 3*b* summarizes schematically the salient features of the primary structure of the Tsr transducer. The aspartate-maltose transducer (Tar) has the same linear arrangement of domains (A. Krikos, personal communication) and a comparison of the Tsr and Tar sequences reveals that the two proteins are 75% identical in residues lying to the carboxy-terminal side of the membrane-spanning peptide, but only 35% identical at the amino-terminal side. This further supports the view that the amino-terminal domain performs the transducer-specific function of chemoeffector binding.

To explain how chemoeffector binding at the periplasmic face of the membrane influences the signalling activity of the transducers at the cytoplasmic face, it is usual to invoke the transmembrane propagation of a conformational change. If the transmembrane structure of Tsr is as simple as we have proposed here, it is not easy to imagine how a single, short peptide chain can perform this function. Studies using cross-linking reagents²⁷ suggest that the transducers may exist in a multimeric form. Thus, transmembrane transmission could occur by the interaction of polypeptide chains in a complex. Alternatively, there may be an equilibrium between monomeric and aggregated forms of the transducers in the membrane. In this model, the unoccupied periplasmic domain could act merely to prevent transducer aggregation. Chemoeffector binding might release this steric block, allowing the formation of an actively signalling transducer multimer.

Several major questions about the function of the transducer proteins remain unanswered: what is the nature of the signal transmitted to the flagellar motors? How does ligand binding at the periplasmic domain influence signalling? How does methylation modulate signalling to produce sensory adaptation? In a more general form these questions apply to signal detecting and processing proteins in other systems. Knowledge of the amino acid sequence and the structure that it implies for the Tsr protein will allow us to design experiments using localized mutagenesis to test critically alternative models for the mechanism of transmembrane signal transmission.

We thank Drs M. Kehry and F. W. Dahlquist for communicating information about Tsr methylated peptides before publication. This work was supported by a USPHS grant.

- Wang, E. A. & Koshland, D. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7157-7161 (1980).
- Boyd, A. & Simon, M. J. *Bact.* **143**, 809-815 (1980).
- Chelsky, D. & Dahlquist, F. W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2434-2438 (1980).
- DeFranco, A. L. & Koshland, D. E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2439-2443 (1980).
- Engstrom, P. & Hazelbauer, G. L. *Cell* **20**, 165-171 (1980).
- Kehry, M. R. & Dahlquist, F. W. *J. biol. Chem.* **257**, 10378-10386 (1982).
- Kehry, M. R. & Dahlquist, F. W. *Cell* **29**, 761-772 (1982).
- Boyd, A., Krikos, A. & Simon, M. *Cell* **26**, 333-343 (1981).
- Parkinson, J. S., Slocum, M. K., Callahan, A. M., Sherris, D. & Houts, S. E. in *Mobility and Recognition in Cell Biology* (eds Sund, H. & Veeger, C.) (de Gruyter, New York, in the press).
- Shine, J. & Dalgarno, J. *Proc. natn. Acad. Sci. U.S.A.* **71**, 7342-7346 (1974).
- Platt, T. *Cell* **24**, 10-23 (1981).
- Segrest, J. P. & Feldman, R. J. *J. molec. Biol.* **87**, 853-858 (1974).
- Emr, S. D., Hall, M. N. & Silhavy, T. J. *J. Cell Biol.* **86**, 701-711 (1980).
- Dunn, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6744-6748 (1981).
- Buchel, D. E., Gronenborn, B. & Muller-Hill, B. *Nature* **283**, 541-545 (1980).
- Higgins, C. F. et al. *Nature* **298**, 723-727 (1982).
- Young, I. G., Rogers, B. L., Campbell, H. D., Jawarowski, A. & Dhaw, D. C. *Eur. J. Biochem.* **116**, 165-170 (1981).
- Engelman, D. N., Henderson, R., McLachlan, A. D. & Wallace, B. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2023-2027 (1980).
- Kreil, G. A. *Rev. Biochem.* **50**, 317-348 (1981).
- Kehry, M. R., Dahlquist, F. W. & Bond, M. W. in *Mobility and Recognition in Cell Biology* (eds Sund, H. & Veeger, C.) (de Gruyter, New York, in the press).
- Sherris, D. & Parkinson, J. S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6051-6055 (1981).
- Rollins, C. & Dahlquist, F. W. *Cell* **25**, 333-340 (1981).
- Chelsky, D. & Dahlquist, F. W. *Biochemistry* **19**, 4633-4639 (1980).
- Auerswald, E.-A., Ludwig, G. & Schaller, H. *Cold Spring Harb. Symp. quant. Biol.* **45**, 107-113 (1981).
- Smith, H. O. & Birnstiel, M. *Nucleic Acids Res.* **3**, 2387-2398 (1981).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309-321 (1981).
- Staden, R. *Nucleic Acids Res.* **8**, 3673-3694 (1980).

Activation of *nif* gene expression in *Azotobacter* by the *nifA* gene product of *Klebsiella pneumoniae*

Christina Kennedy & Robert L. Robson

ARC Unit of Nitrogen Fixation, University of Sussex, Brighton BN1 9RQ, UK

Structural similarity of nitrogenase, the enzyme responsible for biological nitrogen fixation, from various diazotrophic bacteria has been shown by intergeneric biochemical complementation of component proteins *in vitro*¹, DNA and protein sequence analysis²⁻⁷, and DNA hybridization between *nif* (nitrogen fixation) structural genes from *Klebsiella pneumoniae* and genomic sequences from other nitrogen-fixing bacteria^{8,9}. Despite nitrogenase homologies, little is known about similarities among *nif* regulatory mechanisms although repression of nitrogenase synthesis by NH₄⁺ and O₂ occurs in most diazotrophs¹⁰. In *K. pneumoniae*, the *ntr* (*gln*) genes concerned with regulation of nitrogen metabolism control expression of the *nifLA* operon¹¹⁻¹⁶ whose products act as repressor (*nifL*)^{14,17,18} and activator (*nifA*)¹⁹⁻²¹ of the seven other *nif* transcriptional units. Here we report that the *nifA* gene product of *K. pneumoniae* can activate expression of *nif* genes in both *Azotobacter vinelandii* and *Azotobacter chroococcum*, organisms whose aerobic physiology contrasts with that of the facultative *K. pneumoniae*.

A wide host range multicopy plasmid carrying the *nifA* gene expressed from a constitutive promoter was constructed using pKT230 (ref. 22), a hybrid of RSF1010 (a Q incompatibility group plasmid) and pACYC177 (Fig. 1). A *SalI* fragment from pMC71A²¹ carrying the *nifA* gene and flanking regions spanning portions of the *nifB* and *nifL* genes was ligated at the *XhoI* site within the Km^r (kanamycin-resistance) gene of pKT230. Nif⁺Sm^r (streptomycin resistant) transformants were selected in a *nifA*⁻ mutant (UNF 745) of *K. pneumoniae*. Constitutive expression of *nifA* in *K. pneumoniae* allows derepression of *nif* operons in the presence of NH₄⁺ or O₂ (ref. 21). Thus, as expected, nitrogenase activity was not repressed by NH₄⁺ in transformants that carried a 15-kilobase (kb) recombinant plasmid which we named pCK1. Surprisingly, pCK1

Received 27 September; accepted 17 December 1982.

- Rodbell, M. *Nature* **284**, 17-22 (1980).
- Springer, M. S., Goy, M. F. & Adler, J. *Nature* **264**, 577-579 (1979).
- Boyd, A. & Simon, M. A. *Rev. Physiol.* **44**, 501-517 (1982).
- Hedblom, M. L. & Adler, J. *J. Bact.* **144**, 1048-1060 (1980).

Table 1 Effect of *K. pneumoniae nifA* gene on expression of nitrogenase in *Azotobacter*

Strain		No added N source	CH ₃ COONH ₄ added (initial concentration in mM)
<i>A. vinelandii</i>			
UWr	Nif ⁺	56	ND (10)
UWr (pKT230)		40	ND (10)
UWr (pCK1)		54	58 (24)
UW1r	Nif ⁻ (I ⁻ II ⁻)	No growth	ND (10)
UW1r (pCK1)		36	45 (15)
<i>A. chroococcum</i>			
MCD-1	Nif ⁺	47	ND (15)
MCD-1 (pCK1)		39	28 (15)
MCD-1008	Nif ⁻ (I ⁻ II ⁻)	No growth	ND (15)
MCD-1008 (pCK1)		82	51 (15)

The table shows *in vivo* nitrogenase activity (nmol C₂H₂ reduced per mg protein per min). Organisms were grown aerobically at 30 °C with or without added fixed N (as CH₃COONH₄) in liquid Burks' medium⁴⁰ with sucrose at 1.0% w/v as the carbon source. C₂H₂ reduction was measured *in situ* in exponentially growing 10-ml cultures in 50-ml Erlenmeyer flasks. The values shown are the mean of three independent experiments. Assays were started by sealing flasks with suba seals and injecting C₂H₂ to 10% v/v of the gas phase. Vapour phase samples were withdrawn after 2 h and assayed for C₂H₄ as described elsewhere⁴¹. *A. vinelandii* strains UWr and UW1r, resistant to rifampicin, were provided by P. Bishop. *A. chroococcum* strain MCD-1 and the *nif*⁻ mutant MCD1008 (I⁻ II⁻) were derived in this laboratory and will be described fully elsewhere. ND, none detected.

conferred kanamycin resistance (as did pMC73A carrying the *nifA* fragment inserted at the *Xho*I site in pACYC177 (ref. 21) used for comparison) but another recombinant plasmid having the *Sal*I fragment in the opposite orientation did not. As the cloned *Sal*I fragment carries a portion of the *nifB* gene and its promoter (ref. 23 and M. Cannon, personal communication), kanamycin resistance is apparently regulated by the *nifA* gene product. This inference was supported by observation of decreased kanamycin resistance, estimated from the zone of inhibition around kanamycin disks, in UNF745 (pCK1) with increasing growth temperature and concentration of supplied NH₄⁺ or O₂. These factors diminish expression of *nif* promoters other than *nifLA* in a *nifL*⁺ background due to temperature sensitivity of the *nifL*¹⁷ and/or *nifA*²¹ gene products and to *nifL*-mediated repression in response to NH₄⁺ and O₂ (ref. 17). Also, pCK1 failed to confer kanamycin resistance in a strain of *Escherichia coli* carrying a mutation in *ntrA* (*glnF*), a gene necessary for *nifA* product activity²⁴⁻²⁶.

The *Xho*I site is 30 base pairs (bp) downstream from the initiation codon of the Km^r gene in Tn903 (ref. 27) contained in pKT230 and pACYC177 (refs 22, 28). Thus, ligation at this site with the *nifA* *Sal*I fragment should result in a translational fusion with *nifB*. Expression of a hybrid protein which retains kanamycin phosphotransferase activity in pCK1 and pMC73A is therefore from the *nifA*-dependent *nifB* promoter.

Neither pKT230 nor pCK1a are self-transmissible but they were mobilized by the Tra⁺ plasmid, pRK2013 (ref. 29). Transfer from *E. coli* JC5466 (pCK1) (pRK2013) to *A. vinelandii* (UWr) occurred at a frequency of 10⁻¹–10⁻² per recipient and to *A. chroococcum* (MCD1) at 10⁻³–10⁻⁴. Both UWr (pCK1) and MCD1 (pCK1) had nitrogenase activity when grown with high NH₄⁺ concentrations (Table 1). Also, if transconjugants were grown in the absence of fixed nitrogen, the addition of either NH₄⁺ or normally repressive amounts of O₂ (ref. 30) did not inhibit synthesis (Fig. 2). Thus, repression by both major effectors of *nif* expression, oxygen and ammonia, does not occur if the *K. pneumoniae nifA* gene is expressed constitutively in *Azotobacter*.

The pattern of kanamycin resistance observed in the *Azotobacter* strains carrying pCK1 was similar to that of *K.*

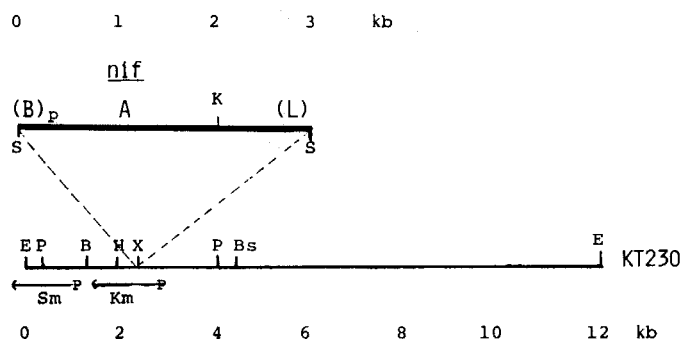


Fig. 1 Restriction map of pCK1: E, *Eco*RI; P, *Pst*I; B, *Bal*I; H, *Hind*II; X, *Xho*I; Bs, *Bst*EII; K, *Kpn*I; S, *Sal*I. Restriction sites shown were as reported^{22,23} and affirmed during this work (the *Kpn*I site in *nifA* was found by V. Buchanan-Wollaston, personal communication). p, Promoter sites. (B) and (L) indicate that incomplete portions of the *nifB* and *nifL* genes, respectively, are contained in the *Sal*I fragment.

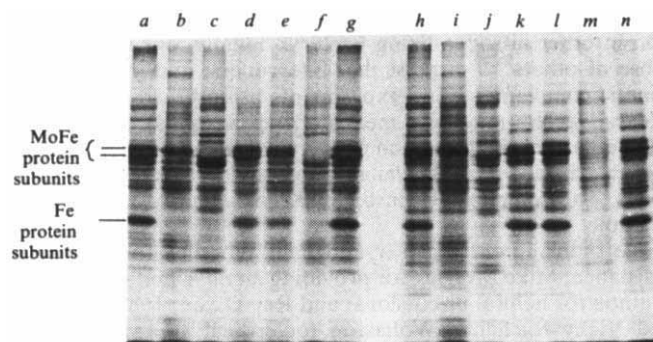


Fig. 2 Influence of pCK1 on effect of NH₄⁺ and O₂ on nitrogenase synthesis in *A. vinelandii* and *A. chroococcum*. Wild-type organisms were grown aerobically at 30 °C in liquid Burks' medium⁴⁰ containing no added fixed N source but modified to contain SO₄²⁻ initially at 50 μM. Nif⁻ mutants were grown with urea (10 mM) as nitrogen source, but were resuspended in the nitrogen-free medium before pulse-labelling; 1-ml portions of cultures were pulse-labelled for 1 h with 10 μCi of ³⁵S₄, and radioactively labelled polypeptides were analysed by SDS-polyacrylamide gel electrophoresis as described elsewhere³⁰. NH₄⁺ repression was tested by adding CH₃COONH₄ to 15 mM (+NH₄⁺) and O₂ repression by increasing the P_{O2} in the gas phase above the cultures to 50% (high O₂) 14 min before addition of isotope. *A. vinelandii* UWr: a, -NH₄⁺ in air; b, +NH₄⁺ in air; c, -NH₄⁺ in high O₂. UWr (pCK1): d, +NH₄⁺ in air; e, -NH₄⁺ in high O₂. UW1r (Nif⁻ I⁻ II⁻): f, -NH₄⁺ in air. UW1r (pCK1): g, -NH₄⁺ in air. *A. chroococcum* MCD-1: h, -NH₄⁺ in air; i, +NH₄⁺ in air; j, -NH₄⁺ in high O₂. MCD-1 (pCK1): k, +NH₄⁺ in air; l, -NH₄⁺ in high O₂. MCD-1008 (Nif⁻ I⁻ II⁻): m, -NH₄⁺ in air. MCD-1008 (pCK1): n, -NH₄⁺ in air.

pneumoniae. Resistance decreased with increasing temperature and NH₄⁺ concentration (data not shown). pCK1 also decreased the growth of both *Azotobacter* species to 50–80% of the normal rate irrespective of nitrogen source. As the *nifA* gene product activates expression from other promoters under *ntrC* control in *K. pneumoniae*^{25,26}, multicopies of *nifA* in *Azotobacter* may perturb the balance of *ntr*-related enzymes.

Nif⁻ mutants of *Azotobacter* suspected to be of a regulatory type are those lacking both nitrogenase components (I⁻ II⁻). In one such mutant of *A. vinelandii* (UW1)³¹, and in five independently isolated I⁻ II⁻ mutants of *A. chroococcum*, a Nif⁺ phenotype was restored by pCK1. Nitrogenase activity was similar to wild-type, was expressed constitutively in NH₄⁺ (Table 1) and was not repressed by high O₂ levels (Fig. 2). As expected, transfer of pCK1 to *A. vinelandii* UW91, a I⁻ II⁻ mutant³², or to *A. chroococcum* MCD1002 (I⁻ II⁻) or MCD1004 (I⁻ II⁻) mutants gave no Nif⁺ transconjugants. Elimination of pCK1 (Sm^rKm^r) from UW1r was achieved by introducing pKT248, a Sm^rCm^r (chloramphenicol-resistant) IncQ plasmid²²; all of 48 Cm^rKm^r transconjugants were Nif⁻, showing that the Nif⁺

phenotype of the UW1r (pCK1) transconjugants was due to a *trans* effect of the *nifA* gene product encoded by the plasmid. Similar results were observed for MCD1008 (pCK1).

Our results suggest that a protein analogous to the *nifA* product activates expression of other *nif* genes in *Azotobacter* as it does in *K. pneumoniae*. An earlier study concluded that the mechanism for ammonia repression of nitrogenase is similar in both organisms³³. Evidence that species of *Rhizobium* may also contain a *nifA* gene was indicated by the correlation in cowpea varieties of *ex planta* nitrogenase activity with production of a pink-coloured compound in the presence of 6-cyanopurine³⁴, an activity associated with *nifA* function in *K. pneumoniae*³⁵. Also, it has been shown recently that the *nifA* product of *K. pneumoniae* can activate expression of the *Rhizobium meliloti nifH* promoter cloned in *E. coli*³⁶ and that *nifA* DNA hybridizes to sequences contained in the *Fix* (*nif*) region of a nodulation plasmid from *Rhizobium leguminosarum*³⁷. Thus, it seems likely that *nifA* activation is ubiquitous among diazotrophs and it will be of great interest to determine whether constitutive production of the *K. pneumoniae nifA* product results in constitutive nitrogenase production in other nitrogen-fixing bacteria. Finally, our results and those of others^{38,39} suggest that intergeneric and interfamilial genetic complementation experiments using broad host range cloning vectors may be a means of bypassing more laborious biochemical characterization of mutants when at least one well characterized system is available to provide cloned reference genes. The complex *nif* cluster of *K. pneumoniae* is one such example.

We thank Linda Batsleer and Paul Woodley for technical assistance; Mike Merrick for providing an *ntfA* mutant; Maura Cannon for helpful discussions; and Ray Dixon, John Postgate and Vicky Buchanan-Wollaston for critical appraisal of the manuscript.

Received 8 October; accepted 5 January 1983.

- Emerich, D. W., Hageman, R. V. & Burris, R. H. *Adv. Enzym.* **52**, 1–22 (1981).
- Hausinger, R. P. & Howard, J. B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3826–3830 (1980).
- Mevarech, M., Rice, D. & Haselkorn, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6476–6489 (1980).
- Hase, T., Nakano, T., Matsubara, H. & Zumft, W. G. *J. Biochem.* **90**, 295–298 (1981).
- Sundaresan, V. & Ausubel, F. M. *J. biol. Chem.* **256**, 2808–2812 (1981).
- Scott, K. F., Rolfe, B. G. & Shine, J. J. *J. molec. appl. Genet.* **1**, 71–81 (1981).
- Török, I. & Kondorosi, A. *Nucleic Acids Res.* **9**, 5711–5723 (1981).
- Ruvkun, G. B. & Ausubel, F. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 191–195 (1980).
- Mazur, B. J., Rice, D. & Haselkorn, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 186–190 (1980).
- Eady, R. in *Current Perspectives in Nitrogen Fixation* (eds Gibson, A. H. & Newton, W.) 172–181 (Australian Academy of Science, Canberra, 1981).
- MacNeil, D. & Brill, W. J. *J. Bact.* **144**, 744–751 (1980).
- Espin, G., Alvarez-Morales, A. & Merrick, M. *Molec. gen. Genet.* **184**, 213–217 (1981).
- Sibold, L., Melick, D. & Elmerich, C. *FEMS microbiol. Lett.* **10**, 37–41 (1981).
- Hill, S., Kennedy, C., Kavanagh, E., Goldberg, R. B. & Hanau, R. *Nature* **290**, 424–426 (1981).
- de Bruijn, F. J. & Ausubel, F. M. *Molec. gen. Genet.* **183**, 289–297 (1981).
- Espin, G., Alvarez-Morales, A., Cannon, F., Dixon, R. & Merrick, M. *Molec. gen. Genet.* **186**, 518–524 (1982).
- Merrick, M. *et al. Molec. gen. Genet.* **185**, 75–81 (1982).
- Buchanan-Wollaston, V., Cannon, M. C. & Cannon, F. C. *Molec. gen. Genet.* **184**, 102–106 (1981).
- Roberts, G. P., MacNeil, T., MacNeil, D. & Brill, W. J. *J. Bact.* **136**, 267–279 (1978).
- Dixon, R. *et al. Nature* **286**, 128–132 (1980).
- Buchanan-Wollaston, V., Cannon, M. C., Beynon, J. L. & Cannon, F. C. *Nature* **294**, 776–778 (1981).
- Bagdasarian, M. *et al. Gene* **16**, 237–247 (1981).
- Riedel, G. E., Ausubel, F. M. & Cannon, F. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2866–2870 (1979).
- Sibold, L. & Elmerich, C. *EMBO J.* **1**, 1551–1558 (1982).
- Ow, D. W. & Ausubel, F. M. *Nature* **301**, 307–313 (1983).
- Merrick, M. *EMBO J.* (in the press).
- Grindley, N. D. F. & Joyce, C. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7176–7180 (1980).
- Chang, A. C. Y. & Cohen, N. J. *Bact.* **134**, 1141–1156 (1978).
- Figurski, D. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1648–1652 (1979).
- Robson, R. L. *FEMS microbiol. Lett.* **5**, 259–262 (1979).
- Fisher, R. J. & Brill, W. J. *Biochim. biophys. Acta* **184**, 99–105 (1969).
- Shah, V. K., Davis, L. C., Gordan, J. K., Orme-Johnson, W. H. & Brill, W. J. *Biochim. biophys. Acta* **292**, 246–255 (1973).
- Cannon, F. C. & Postgate, J. R. *Nature* **260**, 271–272 (1976).
- Kennedy, C., Dreyfus, B. & Brockwell, J. J. *gen. Microbiol.* **125**, 233–240 (1981).
- MacNeil, D. & Brill, W. J. *J. Bact.* **136**, 247–252 (1978).
- Sundaresan, V., Jones, J., Ow, D. & Ausubel, F. *Nature* (in the press).
- Downie, J. A., Hombrecher, G., Ma, Q.-S., Knight, C. D. & Johnston, A. W. B. in *Molecular Genetics of the Bacterial-Plant Association* (ed. Pühler, A.) (Springer, Berlin, in the press).
- Alikhanian, S. I., Gribanova, L. K. & Mishina, I. N. *Gene* **11**, 177–179 (1980).
- Fischer, R., Tuli, R. & Haselkorn, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3393–3397 (1981).
- Newton, J. W., Wilson, P. W. & Burris, R. H. *J. biol. Chem.* **204**, 445–451 (1953).
- Dalton, H. & Postgate, J. R. *J. gen. Microbiol.* **56**, 307–319 (1969).

Structure of a genomic clone encoding biologically active human relaxin

P. Hudson, J. Haley, M. John, M. Cronk*, R. Crawford, J. Haralambidis, G. Tregear, J. Shine* & H. Niall

Howard Florey Institute of Experimental Physiology and Medicine, University of Melbourne, Parkville, Victoria, Australia 3052

* Centre for Recombinant DNA Research, Research School of Biological Sciences, Australian National University, Canberra, ACT, Australia 2601

Relaxin is a peptide hormone^{1–4} synthesized in the corpora lutea of ovaries during pregnancy and is released into the blood stream prior to parturition. Its major biological effect is to remodel the mammalian reproductive tract to facilitate the birth process⁵. Determination of the structure of human relaxin is thus a first step in opening up the possibility of clinical intervention in cases of difficult labour. However, the limited availability of human ovaries during pregnancy has prevented both direct amino acid sequence determination and isolation of cDNA clones obtained from relaxin producing tissue. Our approach has therefore been to screen directly for a human relaxin gene using an homologous porcine relaxin cDNA probe. We report here the successful identification of a genomic clone from which the structure of the entire coding region of a human preprorelaxin gene has been determined. Synthesis of biologically active relaxin has shown that the novel gene structure described herein codes for an authentic human relaxin. We believe this is the first successful synthesis of a biologically active hormone whose structure was predicted solely from the structure of a genomic clone.

The structural genes for both rat⁶ and pig⁷ relaxins have been identified in cDNA libraries constructed from the mRNA of ovarian tissue. Both relaxin genes were found to code for a single chain precursor which resembled preproinsulin⁸ in the overall configuration: (signal peptide/B chain/C peptide/A chain). The biologically active hormone in all species examined so far consists of two peptide chains (A and B chains) linked by one intra-chain and two inter-chain disulphide bonds.

Porcine and rat preprorelaxins contain unexpectedly large connecting (C) peptides of 104 and 105 residues respectively in comparison to porcine insulin with a C peptide of about 30 residues⁸. When comparing rat and pig relaxins some regions of the C peptide were found to possess a higher degree of nucleotide sequence homology than the A and B chains⁷. Hence a specific probe was made from a short (150 base pair) fragment of the porcine relaxin cDNA clone corresponding to amino acids 45–94 in the C peptide⁷ which was the region of maximum homology (71% at the nucleotide level) between rat and porcine relaxin sequences. Screening a human genomic library^{9,10} with this porcine relaxin cDNA probe yielded two λ clones (λ H5 and λ H7) which hybridized strongly and were therefore likely to contain coding sequences of human relaxin.

The human genomic clones were further characterized by restriction enzyme analysis using as probes two separate fragments of porcine relaxin cDNA; one containing the porcine A chain cDNA sequence and the other encoding the B chain sequence. The two fragments were generated by cleavage of the porcine relaxin cDNA clone at a single *HpaII* site which corresponds (within a few bases) to an intron site in the homologous rat relaxin gene (ref. 6 and our unpublished data). Southern blot analysis¹¹ of the λ H5 and λ H7 clones subsequently revealed that the B and A chain coding regions of the human relaxin gene are positioned on exons I and II respectively (Fig. 1) and these are separated by a single intron of 3.7 kilobases (kb). The strategy used for the complete sequence analysis of exons I and II was to subclone restriction digests of the λ H5 and λ H7 genomic clones into M13 bacteriophage vectors and then screen recombinant phage using

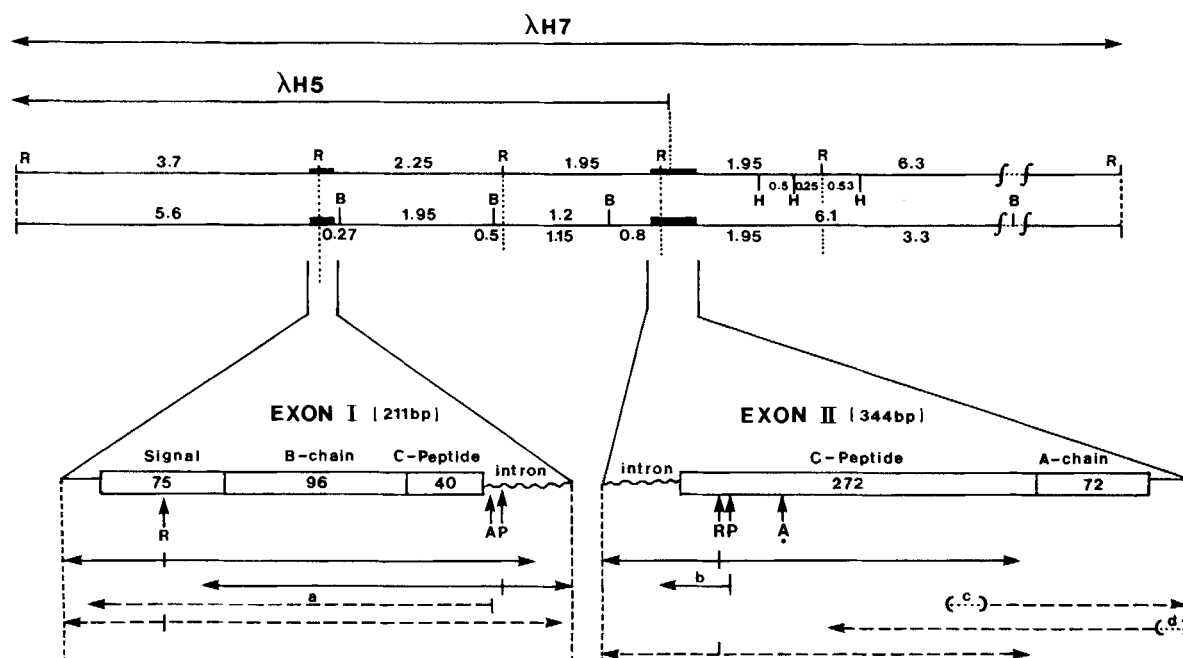


Fig. 1 An abbreviated restriction enzyme map and sequencing strategy on two genomic clones (λ H5 and λ H7) which hybridized with the porcine relaxin cDNA probe. Radiolabelled probes were prepared by primed synthesis on various DNA fragments (100–200 ng) after denaturation with random primers of calf thymus DNA²⁸. Synthesis was initiated by the addition of a 30 μ l reaction mixture containing 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM dithiothreitol, 10 mM MgCl₂, 5 units of *Escherichia coli* DNA polymerase 1 (Klenow fragment), 500 μ M each of dCTP, dGTP, dTTP and 0.3 μ M [α -³²P]dATP ($\sim 3,000$ Ci mmol⁻¹ Amersham). After incubation at 37°C for 30 min the reaction was terminated by dilution into 300 μ l of a buffer containing 0.3 M NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA and passed through a Sephadex-G50 column (1 cm $\times$ 5 cm) in the same buffer. The radiolabelled probe was collected from the peak fractions at void volume and precipitated with 2 volumes of ethanol at -20°C for 2 h using tRNA (10 μ g) as carrier. λ Phage containing genomic DNA fragments⁹ were grown on *E. coli* LE392 in soft agar at $\sim 10^5$ phage per 13-cm diameter plate and transferred to nitrocellulose filters (Shleicher and Schull BA85) as described by Benton and Davis²⁹. Filters were hybridized with the radiolabelled probe at 40°C for 18 h in modified Denhardt's solution³⁰ containing 5 $\times$ SSC and 25% formamide. Filters were washed in 2 $\times$ SSC at 30°C for 1 h before exposing to X-ray film (Kodak XS-5) for 24 h. Regions of the plate which exhibited positive hybridization were subcultured and rescreened successively until single positive plaques could be selected. Phage were collected after lysis of 1-litre cultures of *E. coli* DP50supF cells and DNA prepared by the methods described by Maniatis *et al.*¹⁰ and Yamamoto *et al.*³¹. For DNA sequence analysis restriction fragments of the recombinant phage were subcloned directly by ligation into the *Eco*RI, *Pst*I or *Sma*I site of the M13 vectors mp 7.1, mp8 or mp9 followed by transformation into *E. coli* JM101 cells³². Plaques containing the coding regions were selected by a similar technique as described for the genomic screens above, except the M13 phage were plated at lower density (10^3 phage per 9-cm plate). Positive plaques were grown for a preparative yield of either single-stranded template or double-stranded (rf) form³². Single-stranded templates were sequenced directly by the primed synthesis method of Sanger *et al.*³³ (shown as broken lines) using either an M13-specific primer (Collaborative Research) or synthetic primers complementary to various sequences in the coding region. Further sequence analysis of the subclones was obtained by directly end-labelling fragments and sequencing by the chemical degradation method of Maxam and Gilbert³⁴ (solid lines). Sizes are given in kilobase-pairs (kb) and cleavage sites are designated *Eco*RI (R), *Bgl*II (B), *Hind*III (H), *Hpa*II (P) and *Alu*I (A). The genomic clone λ H5 terminates at an *Eco*RI linker attached to the *Alu*I site (Δ) in the C peptide (Exon II). The definitive nucleotide sequence over the coding region was compiled from the genomic clone λ H7 by subcloning *Eco*RI fragments into M13mp8. Overlapping sequences which confirm the *Eco*RI sites were obtained from runs a, b on subcloned *Alu*I and *Pst*I fragments of λ H7 respectively. The primers used for sequencing were c, 5'TTCGCAAT-AGGCA, and d, 5'GCACAATTAGCT, prepared using the phosphite chemistry method described by Beaucage and Caruthers³⁵.

porcine relaxin probes specific for either exon region. Positive subclones were sequenced by the techniques summarized in Fig. 1.

The amino acid sequence of human preprorelaxin as deduced from the nucleotide sequence of the λ H7 genomic clone is presented in Fig. 2. Positions of the initiation and termination codons and hence the reading frame were unambiguous from the homology to porcine relaxin. In the C peptide there is only one position for intron excision which retains the reading frame homology with porcine and rat relaxins and the consensus sequences for exon/intron/exon boundaries¹². Human relaxin and human insulin genes contain¹³ an intron at a similar position in the C peptide supporting the suggestion that these hormones arose by a gene duplication. If this is true, subsequent evolutionary divergence of insulin and relaxin has resulted in only limited sequence homology in the A and B chains (Fig. 3) and no significant homology in the signal and connecting peptides.

Interpretation of the processing sites in human preprorelaxin relies on a comparison to the homologous structures determined for porcine and rat relaxins (Fig. 3). Cleavage of the signal peptide, presumably by a membrane peptidase after a short side-chain amino acid in agreement with the consensus of signal processing¹⁴, should occur at either Ala -1, -2 or -4 or at Ser

-6. Homology with porcine relaxin would favour cleavage at Ala -1 removing a signal peptide of 25 amino acids (Fig. 2). This would be analogous to human preproinsulin in which the signal peptide is cleaved immediately following three adjacent alanines (that is, positions -1, -2, -3). For synthetic studies of human relaxin knowledge of the exact cleavage position is not required since it can be predicted from structure-activity studies on synthetic porcine relaxin¹⁵ and from comparative models of the tertiary structure^{16,17} that the core sequence of the human relaxin B chain (amino acids Lys 1 to Ser 25) contains all the essential elements for biological activity.

Although both the A and B chains of human relaxin contain dibasic residues we propose that excision of a 104 residue C peptide occurs in a similar manner as in porcine and rat relaxin. Cleavage at the B-chain/C-peptide junction almost certainly occurs at Leu 32 since this is observed in both porcine and rat relaxins. Cleavage at the C-peptide/A-chain junction probably occurs at the group of 4 basic residues by homology to rat relaxin. However, it is likely that the Arg-Pro imide bond would be resistant to proteolysis since neither endopeptidases like trypsin¹⁸ nor aminopeptidases¹⁹ readily cleave this peptide linkage. Thus the A chain of human relaxin probably contains Arg-Pro (residues 137–138 of preprorelaxin) at the amino-

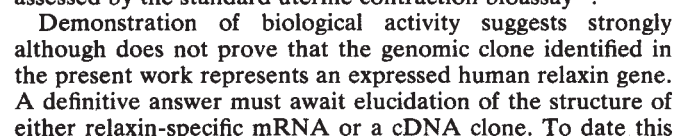


Fig. 3 Comparison of the amino acid sequences of the B and A chains between human relaxin, human insulin⁸ and the relaxin family. Residues boxed in human insulin are identical to human relaxin. Boxed areas in the relaxin series highlight similarities to human and porcine relaxins. The numbering system follows that of Fig. 2, assuming Lys 1 is the amino terminal residue of the B chain. Arrows indicate probable sites of proteolytic cleavage with confirmation of the amino terminal residue of the B and A chains of porcine^{1,2} rat³ and shark⁴ relaxins from protein sequencing data. In support of the proposed similarity between the tertiary structures of porcine relaxin and insulin^{16,17} all six cysteine residues in the A and B chains are retained in human relaxin, thus maintaining the overall disulphide bond configuration.



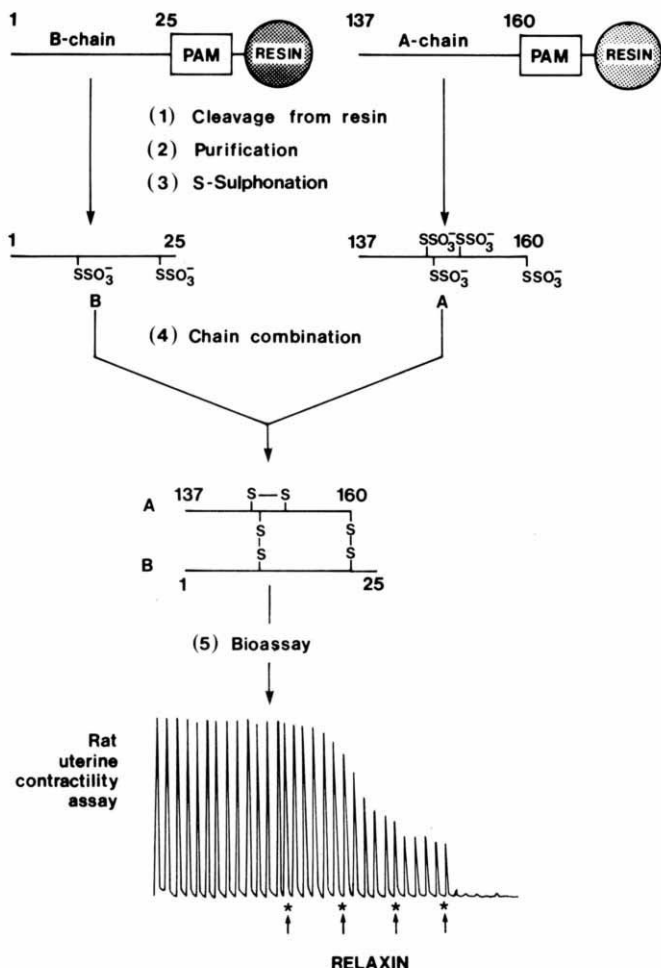


Fig. 4 Chemical synthesis of a peptide with relaxin bioactivity based upon the putative human relaxin sequence deduced from a genomic clone. The A chain (Arg 137-Cys 160) and the B chain segment (Lys 1-Ser 25) were assembled separately on a phenylacetamidomethyl (PAM) polystyrene resin support using procedures previously described¹⁵. Following cleavage from the resin and purification the peptides were S-sulphonated and combined using the insulin chain combination conditions described by Chance and Hoffmann³⁶. Specific relaxin-like biological activity was detected in the rat uterine contractility assay²¹ by adding increasing concentrations of the combination mixture at the positions marked by asterisks. In preliminary experiments combination yields and specific activities relative to porcine relaxin of the order of 1 to 3% were achieved. In control experiments the synthetic A and B chain preparations were each separately exposed to the conditions for chain combination and assayed. There was no inhibition of uterine contractility indicating the absence of non-specific toxic effects on the rat myometrium.

for the bioassay data. Dr T. Maniatis kindly provided the human genome library and R. Staden the computer programs for DNA sequence analysis. The research was carried out under NIH/ASCORD containment guidelines for recombinant DNA techniques and was supported by grants from the National Health and Medical Research Council of Australia, the Myer Family Trusts, the Ian Potter Foundation, the Howard Florey Biomedical Foundation and by NIH grant HD11908.

Received 20 September; accepted 15 December 1982.

- Schwabe, C., McDonald, J. K. & Steinetz, B. G. *Biochem. biophys. Res. Commun.* **75**, 503-510 (1977).
- James, R., Niall, H., Kwok, S. & Bryant-Greenwood, G. *Nature* **267**, 544-546 (1977).
- John, M. J., Borjesson, B. W., Walsh, J. R. & Niall, H. D. *Endocrinology* **108**, 726-729 (1981).
- Schwabe, C., Gowan, L. K. & Reinig, J. W. *Ann. N.Y. Acad. Sci.* **380**, 6-12 (1982).
- Porter, D. G. *Biol. Reprod.* **7**, 458-465 (1972).
- Hudson, P., Haley, J., Cronk, M., Shine, J. & Niall, H. *Nature* **291**, 127-131 (1981).
- Haley, J. *et al. DNA* **1**, 155-162 (1982).
- Chan, S. J., Keim, P. & Steiner, D. F. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1964-1968 (1976).
- Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157-1174 (1978).
- Maniatis, T. *et al. Cell* **15**, 687-701 (1978).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Mount, S. M. *Nucleic Acids Res.* **10**, 459-472 (1982).
- Bell, G. I. *et al. Nature* **284**, 26-32 (1980).
- Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. *Ann. N.Y. Acad. Sci.* **343**, 1-16 (1980).
- Tregear, G. *et al. in Peptides: Synthesis, Structure, Function* (eds Rich, D. H. & Gross, E.) 249-252 (Pierce Chemical Co., Rockville, Illinois, 1981).
- Bedarkar, S., Turnell, W. G., Blundell, T. L. & Schwabe, C. *Nature* **270**, 449-451 (1977).
- Isaacs, N. *et al. Nature* **271**, 278-281 (1978).
- Hirs, C. H. W., Moore, S. & Stein, W. H. *J. biol. Chem.* **219**, 623-630 (1956).
- Light, A. *Meth. Enzym.* **11**, 426-436 (1967).
- Du, Y.-C., Minasian, E., Tregear, G. W. & Leach, S. J. *Int. J. Peptide Protein Res.* **20**, 47-54 (1982).
- Wqvist, N. & Paul, K. G. *Acta endocr.* **29**, 135-136 (1958).
- Bryant-Greenwood, G. D., Jeffrey, R., Ralph, M. M. & Seemark, R. F. *Biol. Reprod.* **23**, 792-798 (1980).
- Weib, M., Magelschmidt, M. & Struck, H. *Horm. Metab. Res.* **11**, 408-410 (1979).
- Stone, M. L., Sedlis, A. & Zuckerman, M. B. *Ann. N.Y. Acad. Sci.* **75**, 1011-1015 (1959).
- Chihai, H. J. & Epsey, L. L. *Endocrinology* **93**, 1441-1445 (1973).
- Loumaye, E., Decoan, S. & Thomas, K. J. *Clin. Endocr. Metab.* **50**, 1142-1143 (1980).
- Hisaw, F. L., Hisaw, F. L. Jr & Alden, D. B. *Endocrinology* **81**, 375-385 (1967).
- Taylor, J. M., Illmensee, R. & Summers, J. *Biochim. biophys. Acta* **442**, 324-330 (1976).
- Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
- Denhardt, D. T. *Biochem. biophys. Res. Commun.* **23**, 641-646 (1966).
- Yamamoto, K. R., Alberts, B. M., Benzinger, R., Lawhorne, L. & Treiber, G. *Virology* **40**, 734-744 (1970).
- Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. *J. molec. Biol.* **143**, 161-178 (1980).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- Beaucage, S. L. & Caruthers, M. H. *Tetrahedron Lett.* **22**, 1859-1862 (1981).
- Chance, R. E. & Hoffmann, J. A. Australian Patent Appl. No. 68844/81.

Mitochondrial DNA sequences in the nuclear genome of a locust

Gerd Gellissen*, J. Y. Bradfield, B. N. White & G. R. Wyatt

Department of Biology, Queen's University, Kingston, Ontario, Canada K7L 3N6

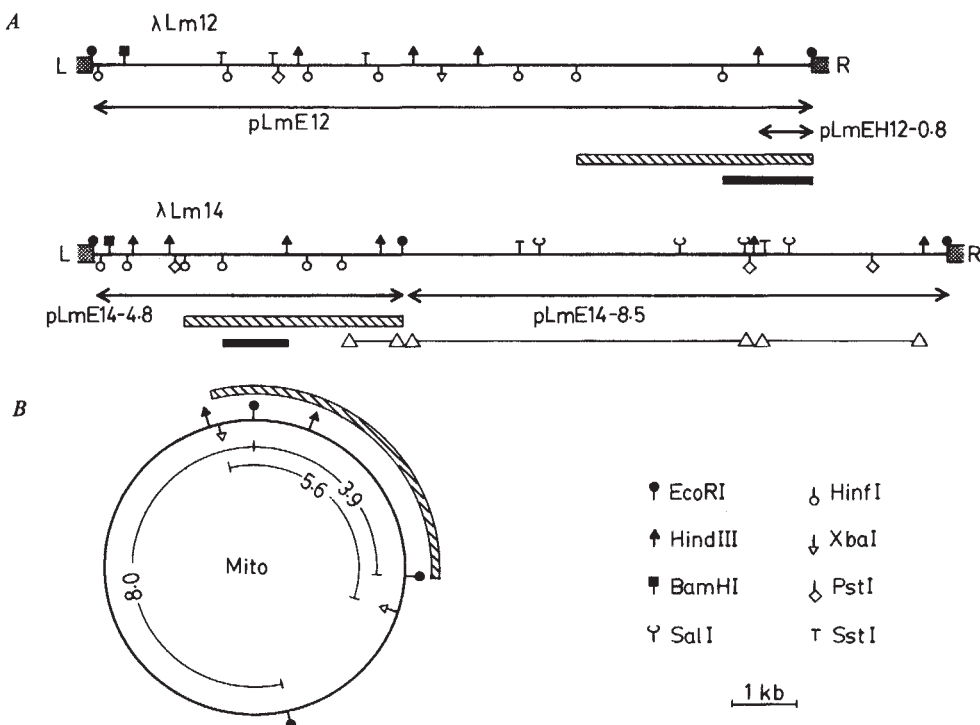
The endosymbiotic theory¹ of the origin of mitochondria is widely accepted, and implies that loss of genes from the mitochondria to the nucleus of eukaryotic cells has occurred over evolutionary time^{2,3}. However, evidence at the DNA sequence level for gene transfer between these organelles has so far been limited to a single example, the demonstration that a mitochondrial ATPase subunit gene of *Neurospora crassa* has an homologous partner in the nuclear genome⁴. From a gene library of the insect, *Locusta migratoria*, we have now isolated two clones, representing separate fragments of nuclear DNA, which contain sequences homologous to the mitochondrial genes for ribosomal RNA, as well as regions of homology with highly repeated nuclear sequences. The results suggest the transfer of sequences between mitochondrial and nuclear genomes, followed by evolutionary divergence.

* Present address: Institut für Zoologie III, Universität Düsseldorf, Universitätsstrasse 1, 4000 Düsseldorf 1, FRG.

has not been possible because of the scarcity of human ovarian tissue during pregnancy and the extremely low abundance of relaxin specific mRNA in the non-pregnant state²². However, restriction fragments of the human relaxin gene available from the present work could be used to screen cDNA clone banks with some hope of finding relaxin-specific cDNA clones in low abundance. In the meantime, synthetic human relaxin can be made available for a variety of physiological and clinical studies bearing on the established functions of this hormone in remodelling connective tissues^{23,24} and in regulating uterine motility^{5,21}. This may lead to the development of clinical treatments to alleviate some of the problems encountered at childbirth. In addition, several more speculative functions of relaxin in ovulation^{22,25}, sperm penetration²⁶ or blastocyst implantation²⁷ can be further investigated.

We thank D. Scanlon, J. Gorman and C. Southwell for assistance with primer synthesis and B. Kemp and S. Denney

Fig. 1 Structures of locust nuclear DNA clones (A) and mitochondrial DNA (B). **A**, nuclear DNA clones. λ Lm12 and λ Lm14 were obtained from a Charon 4 *Eco*RI library of *Locusta* DNA⁵. Subclones in plasmid pAT153, grown in *Escherichia coli* HB101, were: from λ Lm12, the entire insert of 11.3 kb (pLmE12) and a 0.8-kb fragment (pLmEH12-0.8); from λ Lm14, two fragments of 4.8 and 8.5 kb (pLmE14-4.8 and pLmE14-8.5). L and R indicate the left and right arms of λ , respectively. In pLmE14-8.5, the *Hinf*I sites (not yet fully mapped) are omitted. Cross-hatched bars indicate regions that include sequences homologous with mitochondrial DNA (see text). Sequences homologous to the 1.4-kb mitochondrial rRNA (marked by solid bars) were identified by hybridization with cDNA against fat body RNA¹². Lines limited by triangles (Δ — Δ) indicate fragments that include highly repetitive nuclear sequences. **B**, *Locusta* mitochondrial DNA. DNA was isolated from fat body and flight muscle mitochondria¹³ and partially purified by centrifugation in CsCl gradients with ethidium bromide¹⁴, followed by extraction with phenol and chloroform. The cross-hatched bar marks the region that includes sequences homologous with the cloned nuclear DNA sequences; only the restriction sites relevant to localizing this region are shown. *Eco*RI fragments of 8.0, 3.9, 2.3 and 1.4 kb, *Hind*III fragments of 14.2 and 1.5 kb, *Xba*I fragments of 5.6, 4.4, 2.6, 2.0 and 1.2 kb (see Fig. 3B, panel 1) and several *Hinf*I fragments, but no sites for *Bam*HI or *Sal*I, were found.



In the course of a study of gene expression during development in *L. migratoria*, we prepared a DNA library by partial digestion of locust fat body DNA with *Eco*RI, selection for 18–20-kilobase (kb) fragments, and cloning in λ Charon 4 (ref. 5). This library was screened with cDNAs prepared against RNA from fat body of reproductive females, adult males and fifth instar larvae. A group of six clones showed strong hybridization, indicating homology with prevalent sequences in all three cDNA probes. They also showed some degree of homology with each other when tested by dot hybridization with a plasmid subcloned probe.

Four of these clones, tested on Northern blots of locust tissue RNA, showed homology with two RNA species of 1.4 and 0.8 kb. Two clones, λ Lm12 and λ Lm14, were selected for study, and sequences from them were subcloned in a plasmid (Fig. 1A). Restriction mapping showed quite distinct patterns for λ Lm12 and λ Lm14.

Because 1.4 and 0.8 kb are close to the sizes of *Locusta* mitochondrial ribosomal RNAs⁶, and because a poly(A) sequence present in the larger mitochondrial rRNA of another insect, *Drosophila*⁷, provides for initiation of cDNA with an oligo(dT) primer, we examined the possibility that these clones were of mitochondrial origin. Northern blotting showed that the two RNA components to which pLmE14-4.8 showed homology (Fig. 2A) were also homologous with mitochondrial DNA (Fig. 2B), and are presumably mitochondrial rRNAs. They were present in several locust tissues but especially abundant in flight muscle, which is rich in giant mitochondria. pLmE12 gave the same reaction (not shown), but pLmEH12-0.8 hybridized with the larger rRNA only (Fig. 2C).

The suggestion of mitochondrial origin for the λ clones was not supported, however, by comparison of the restriction maps. Mitochondrial DNA, prepared either from fat body or flight muscle mitochondria, was found to be a circular molecule of ~16 kb, similar to that of other animals (Fig. 1B). The restriction map does not match that of λ Lm12 or λ Lm14. However, as the clones contained sequences homologous to mitochondrial rRNA, they must possess some relationship to the mitochondrial genome.

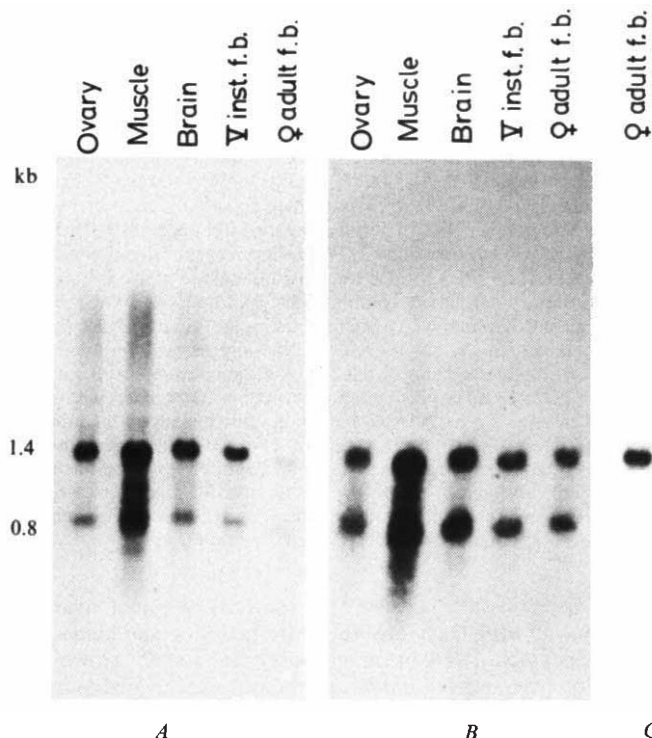


Fig. 2 'Northern blots' showing hybridization of cloned DNA and mitochondrial DNA with RNA from different locust tissues. RNA was extracted¹⁵ from tissues, separated (10 μ g per lane) by electrophoresis on 1.2% agarose gel with methyl mercuric hydroxide¹⁶ and transferred¹⁷ to Gene Screen (NEN). DNA used as probes was labelled by nick translation¹⁸ to specific activity 2.5×10^8 c.p.m. per μ g. RNA samples from the following tissues were used: ovary, flight muscle, brain, fifth instar larval fat body (f.b.), adult female fat body. **A**, hybridization of RNA with pLmE14-4.8 (see Fig. 1). **B**, hybridization of RNA with mitochondrial DNA. **C**, hybridization of RNA with pLmEH12-0.8. Sizes were determined by comparison with *Hind*III fragments of λ DNA.

Fig. 3 Hybridization of locust nuclear (*A*) and mitochondrial (*B*) DNA to cloned probes. *A*, hybridization of nuclear DNA. DNA was prepared¹⁹ from fat body nuclei, digested with *Eco*RI and electrophoresed (15 µg per lane) on 0.8% agarose gel in 40 mM Tris-HCl (pH 8.3), 33 mM NaOH, 1 mM EDTA. Fragments were transferred to diazobenzyloxymethyl paper²⁰ and hybridized (5 × SSC, 50% formamide, 2 × Denhardt's solution, 50 mM phosphate pH 6.5, 42 °C, 12 h, followed by repeated washing in 0.1 × SSC at 65 °C) to cloned DNAs labelled by nick translation. Panel 1, probed with pLmE12 (see Fig. 1); 2, probed with pLmE14-4.8; 3, probed with pLmE14-8.5. *B*, hybridization of partially purified mitochondrial DNA, digested with restriction enzymes indicated. Panel 1, stained gel showing mitochondrial fragments; 2, blot of same hybridized with pLmE12; 3, hybridization with pLmE14-4.8. Size scales: left, λ *Hind*III fragments; right, *Eco*RI fragments of mitochondrial DNA.

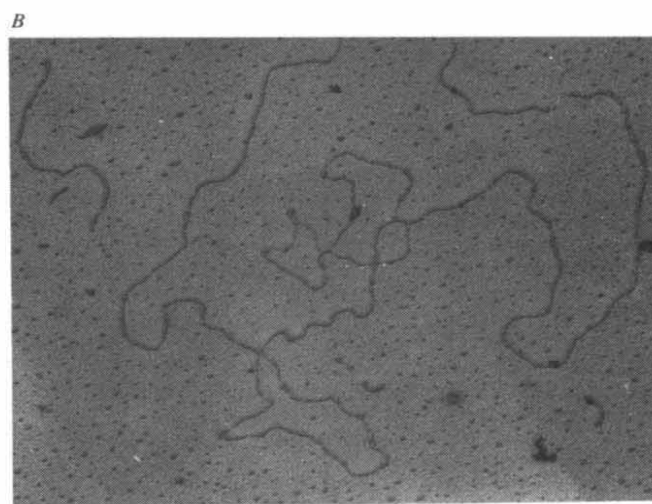
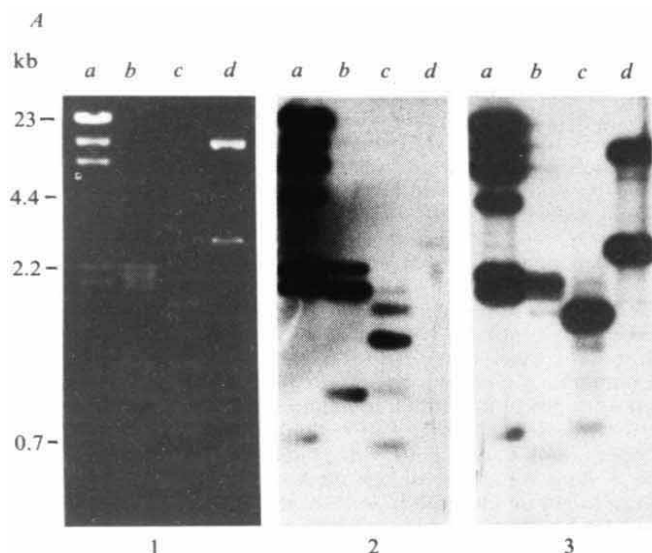
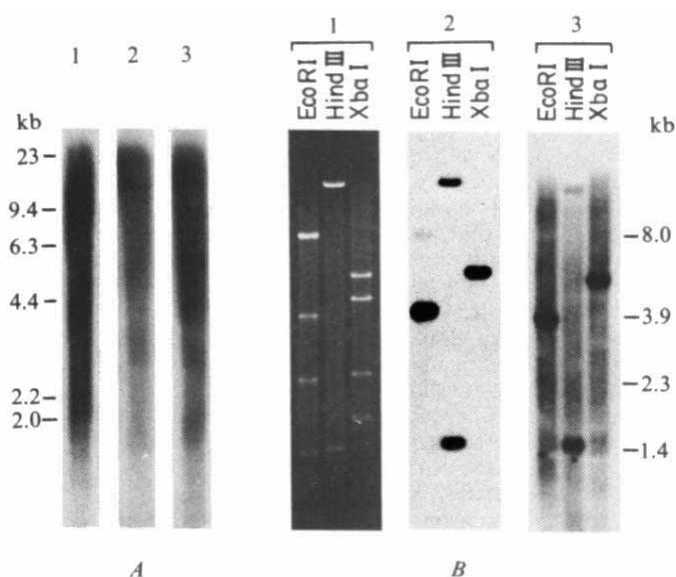
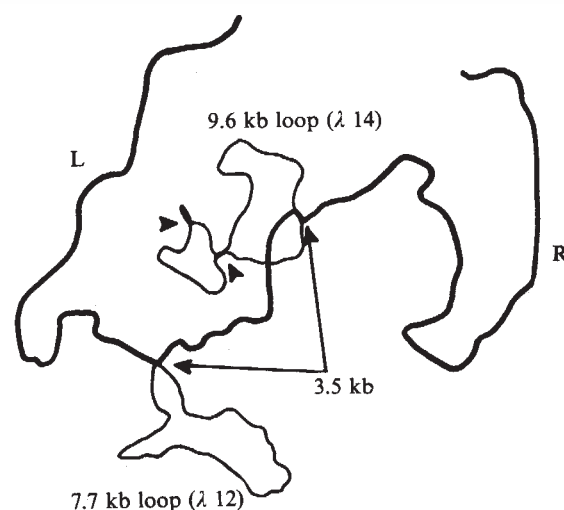


Fig. 4 Homologies between the cloned DNA sequences and nuclear and mitochondrial DNAs. *A*, Southern blots. Digests of cloned DNAs were electrophoresed on 1.5% gel and transferred to nitrocellulose filters¹⁷. Panel 1, gel stained with ethidium bromide; 2, hybridization (3 × SSC, 10 × Denhardt's solution, 65 °C, 12 h, followed by washing as in Fig. 3) of blot with mitochondrial DNA, prepared as described in Fig. 1 legend and labelled by nick translation; 3, hybridization with nuclear DNA, extracted from nuclei from adult female locust fat body²¹, purified in a CsCl ethidium bromide gradient to remove any supercoiled mitochondrial DNA, and labelled by nick translation. Sample *a*, marker λ DNA *Hind*III digest (nick-translated λ DNA was included in the probes); *b*, pLmE12 *Hinf*I digest; *c*, pLmE14-4.8 *Hinf*I digest; *d*, pLmE14-8.5 *Hind*III digest. *B*, heteroduplex analysis^{22,23} showing sequence homology between the clones λ Lm12 and λ Lm14. L, R, left and right arms of the λ Charon 4 vector. Arrows mark snapbacks.



The relationships of the cloned sequences to nuclear and mitochondrial DNA were examined by Southern blotting (Fig. 3). *Eco*RI digests of nuclear DNA probed with pLmE12, pLmE14-4.8 and pLmE14-8.5 (Fig. 3*A*) showed, in each case, a diffuse pattern of hybridization, suggesting homology with highly repeated sequences in the nuclear genome. The hybridization pattern was similar for *Hind*III and *Xba*I digests of nuclear DNA (not shown). In the presence of strong repeated sequence

hybridization, one would not expect to detect hybridization with unique sequence nuclear DNA fragments. Homology between the clones and *Locusta* mitochondrial DNA is examined in Fig. 3*B*. Blots of a preparation enriched for mitochondrial DNA probed with either pLmE12 (Fig. 3*B*, panel 2) or pLmE14-4.8 (Fig. 3*B*, panel 3) showed hybridization to the 3.9-kb *Eco*RI fragment, both *Hind*III fragments and the 5.6-kb *Xba*I fragment. Background hybridization in Fig. 3*B*,

panel 3, is due to contamination of the mitochondrial preparation with nuclear DNA. In addition, pLmE12 hybridized weakly to the 8.0-kb *EcoRI* fragment, suggesting that a small region of homology is shared between the subclone and this mitochondrial fragment. Probing with pLmE14-8.5 (not shown) revealed no homology with mitochondrial DNA. The results indicate that λ Lm12 and λ Lm14 originated from nuclear DNA but contain sequences homologous to portions of the mitochondrial genome. We note that, although restriction mapping shows λ Lm12 and λ Lm14 to represent different sequences from the *Locusta* genome, they show homology with the same restriction fragments of mitochondrial DNA.

The sequences in the clones were further characterized by the experiments shown in Fig. 4. Southern blots of digests of pLmE12 and pLmE14-4.8 showed fragments homologous with mitochondrial DNA, whereas the digest of pLmE14-8.5 did not (Fig. 4A, panel 2; a faint band in lane *d* is probably due to a trace of nuclear DNA in the mitochondrial DNA probe). When probed with nuclear DNA, sequences from pLmE12 hybridized moderately strongly, and one *HinfI* fragment of pLmE14-4.8 and two large *HindIII* fragments of pLmE14-8.5 reacted strongly. This confirms the presence in the clones of sequences homologous with highly repeated nuclear DNA sequences, as suggested by the genomic Southern blots (Fig. 3), and shows that, in λ Lm14, these extend from the right flank of the region of mitochondrial homology. Heteroduplex formation between λ Lm12 and λ Lm14 (Fig. 4B) showed homology between the two clones in the region that bears homology with mitochondrial DNA, even though their restriction patterns differ. The 3.5-kb stretch of heteroduplex formation matches this region. This result was confirmed by Southern hybridization tests between λ Lm12 and pLmE14-4.8 (not shown). Snapback structures within λ Lm14 indicate the presence of inverted repeats, consistent with the presence of highly repeated sequences flanking the regions of homology with mitochondrial DNA.

The data presented demonstrate the presence, in the nuclear genome of a locust, of sequences homologous to mitochondrial DNA, including mitochondrial rRNA genes. The nuclear sequences seem to be moderately repeated, as several cross-hybridizing clones were found. Divergence of the homologous sequences, however, is shown by the lack of correspondence in the restriction maps of two clones and the mitochondrial DNA. The presence of highly repeated sequences in the nuclear clones flanking the mitochondrial homologous regions suggests a role in sequence mobility, as postulated for other mobile DNA elements⁸.

Since this work was done, we have learned that mitochondrial DNA sequences are also represented in the nuclear genome of a sea urchin (H. T. Jacobs, J. W. Posakony and E. H. Davidson, personal communication). Assuming the independent origin of the mitochondrial genome, transposition of sequences from it into the nuclear DNA, followed by divergence, is the most likely interpretation of these observations. The evidence for mobility of DNA elements within genomes⁸ and between species⁹ is thus extended by evidence for mobility between organelles. A report on homology between mitochondrial and chloroplast DNAs in maize^{10,11}, published since submission of this manuscript, leads to similar conclusions.

We thank Elizabeth Belland for RNA from locust tissues, and members of the laboratory for criticism and discussion. The research was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the US NIH (HD-07159). G.G. was the recipient of a fellowship from the Deutsche Forschungsgemeinschaft.

Received 8 October; accepted 15 December 1982.

1. Margulis, L. *Origin of Eukaryotic Cells* (Yale University Press, New Haven, 1970).
2. Wallace, D. C. *Microbiol. Rev.* **46**, 208–240 (1982).
3. Gray, M. W. *Can. J. Biochem.* **60**, 157–171 (1982).
4. van den Boogaart, P., Samallo, J. & Agsteribbe, E. *Nature* **298**, 187–189 (1982).
5. Wyatt, G. R., Locke, J., Bradfield, J. Y., White, B. N. & Deeley, R. G. in *Juvenile Hormone Biochemistry* (eds Pratt, G. E. & Brooks, G. T.) 299–307 (Elsevier, Amsterdam, 1981).
6. Kleinow, W. in *The Biogenesis of Mitochondria* (eds Kroon, C. M. & Saccone, C.) 377–381 (Academic, New York, 1974).
7. Spradling, A., Pardue, M. L. & Penman, S. *J. molec. Biol.* **109**, 559–587 (1977).

8. Shapiro, J. A. & Cordell, B. *Biol. Cell* **43**, 31–54 (1982).
9. Lewin, R. *Science* **217**, 42–43 (1982).
10. Stern, D. B. & Lonsdale, D. M. *Nature* **299**, 698–702 (1982).
11. Ellis, J. *Nature* **299**, 678–679 (1982).
12. Buell, G. N., Wickens, M. P., Payour, F. & Schimke, R. T. *J. biol. Chem.* **253**, 2471–2482 (1978).
13. Kleinow, W. *Hoppe-Seyler's Z. physiol. Chem.* **355**, 1027–1034 (1974).
14. Clewell, D. B. & Helinsky, D. T. *Proc. natn. Acad. Sci. U.S.A.* **62**, 1159–1166 (1969).
15. Chinzei, Y., White, B. N. & Wyatt, G. R. *Can. J. Biochem.* **60**, 243–251 (1982).
16. Bailey, J. M. & Davidson, N. *Analyt. Biochem.* **70**, 75–85 (1976).
17. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
18. Rigby, R. W. S., Dieckman, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
19. Blin, N. & Stafford, D. W. *Nucleic Acids Res.* **3**, 2303–2308 (1976).
20. Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350–5354 (1977).
21. Irvine, D. J. & Brasch, K. *Gen. comp. Endocr.* **45**, 91–99 (1981).
22. Davis, R. W., Simon, M. & Davidson, N. *Meth. Enzym.* **21**, 413–428 (1971).
23. Kidd, S. J. & Gover, D. M. *Cell* **19**, 103–119 (1980).

Deletion mapping of the inducible promoter of human IFN- β gene

Jan Tavernier, Dirk Gheysen, Fred Duerinck, Jose Van der Heyden & Walter Fiers

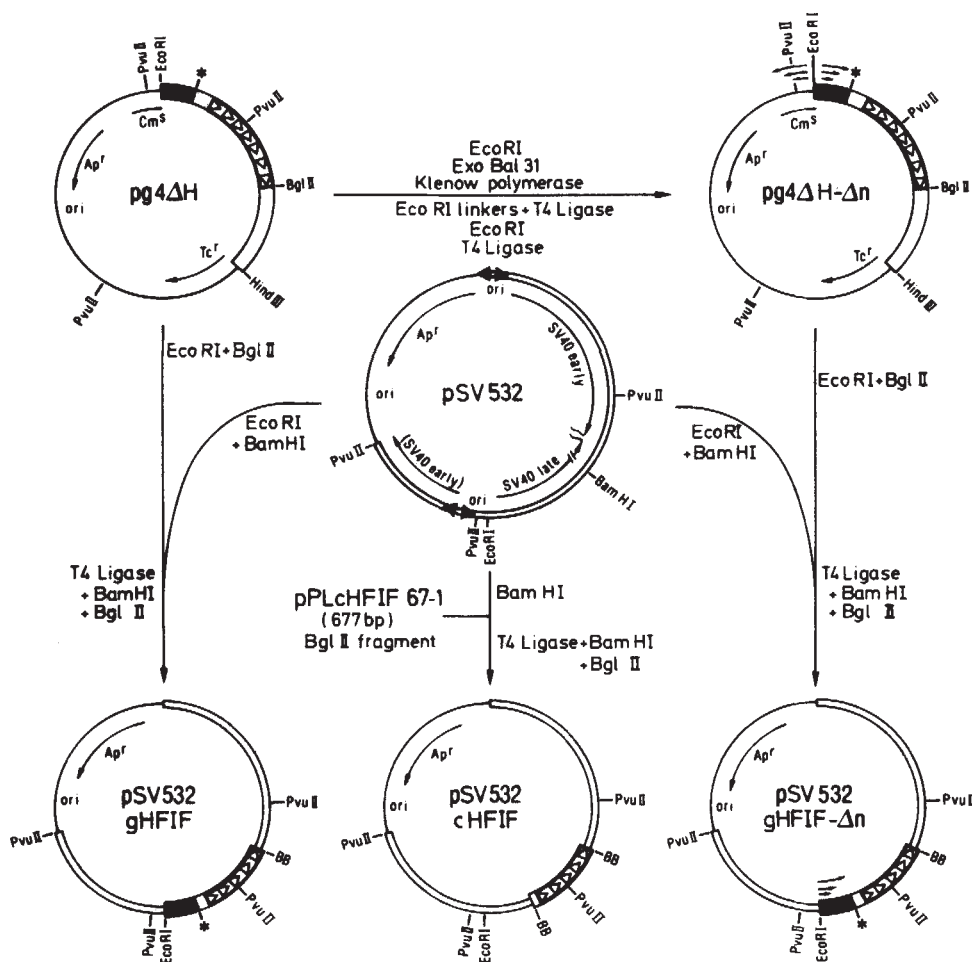
Laboratory of Molecular Biology, State University of Ghent, B-9000 Ghent, Belgium

Human fibroblast cells produce β -type interferon only in response to viral infection or treatment with an inducer such as poly(rI) · poly(rC); this event is most probably controlled at the transcriptional level (for review see ref. 1). To study the induction process, we inserted the human fibroblast interferon (IFN- β) gene, with or without its promoter region, into recombinant simian virus 40 (SV40) plasmid vectors which subsequently were transfected into monkey AP-8 cells. We report here that upon induction with poly(rI) · poly(rC) there was a 10–30-fold increase in IFN- β synthesis. This inducer had no effect on interferon production when the coding region only was inserted into the vector plasmid, which indicates that the promoter region is required for inducibility of this gene. Deletion mapping implicates the region between nucleotides –144 and –186 from the mRNA initiation site in the specific regulation of the IFN- β gene. This region contains a sequence that is remarkably homologous with a consensus sequence found in the 5' flanking region of steroid hormone responsive genes, which might be involved in binding the progesterone-receptor complex².

Interferons are particularly suitable for the study of eukaryotic gene regulation as the interferon system is widespread among eukaryotic organisms and cell types. The very high specific biological activity of interferons renders them easily detectable. Furthermore, the interferon assay allows one to distinguish between different interferons on the basis of their species specificity. Finally, all interferon types, IFN- α (leukocyte), IFN- β (fibroblast) and IFN- γ (immune), have been successfully cloned^{3–10} and their chromosomal genes isolated from human DNA libraries^{11–21}, providing a set of related but differently regulated genes.

To investigate the role of the human IFN- β promoter region on inducibility of the gene by poly(rI) · poly(rC), we used an autonomously replicating SV40 plasmid vector, which has several advantages over the integration of the gene into the genome of the recipient cell. First, it is a rapid method which avoids the time-consuming selection of transformed colonies. Second, the neighbouring area of the cloned gene is defined and well known. Because no integration occurs there is no possibility of unexpected influences of surrounding genomic segments. Finally, variability due to instability of the genome of the transformed cell is avoided. Transfection has, however, the time limitation of a lytic cycle so that long treatments such as superinduction are not possible. Figure 1 shows the recombinant constructions used in the experiments described here. A detailed physical and genetic map of the pSV529 expression vector has been given previously²². It contains 1.4 genome equivalents of SV40 DNA; due to this partial duplication,

Fig. 1 Construction scheme of the chimaeric plasmids pSV532 cHFIF, pSV532 gHFIF and the pSV532 gHFIF- Δn series. pBR322 or pBR325 sequences are shown as a single line, the SV40 sequence as a thin double line and IFN- β sequences as a thick block with arrowheads indicating the coding region and its orientation and a black box indicating the promoter region. In the recombinant plasmid vector pSV532 (centre), a region between the unique *EcoRI* and *BamHI* restriction sites can be replaced by foreign DNA. This expression vector is derived from pSV529 (ref. 22) by a *Bal31* exonuclease-generated deletion of the *EcoRI* and *KpnI* restriction sites (pSV531) and subsequent insertion of an *EcoRI* synthetic linker in the remaining, now unique *KpnI* restriction site of pSV531. The SV40 early and late transcription units are indicated. Arrows show the early and late promoter regions and wavy lines the 3' polyadenylation sites (note that part of the early region is duplicated). pg4 Δ H (upper left) was derived from pg4 HFIF¹³ by deletion of a *HindIII* fragment containing some non-essential vector DNA and part of the 3'-flanking region of the IFN- β gene, leaving a unique *EcoRI* restriction site in front of the gene. pg4 Δ H was restricted with *EcoRI* and subsequently treated with *Bal31* exonuclease for different time intervals to generate progressive deletions in the promoter region. After blunting with Klenow polymerase, *EcoRI* synthetic linkers were inserted (pg4 Δ H- Δn series, with n indicating the deletion site; see also Fig. 2). Only deletions in the promoter region are indicated (series of arrows). The IFN- β gene with its intact promoter region and a series of deletion mutants were then transferred as *EcoRI*-*BglII* fragments into *EcoRI*-*BamHI*-opened pSV532. We counterselected the original constructions by adding both *BamHI* and *BglII* restriction enzymes to the ligation mixture; the recombinant plasmids are referred to as pSV532 gHFIF and pSV532 gHFIF- Δn , respectively. Also, the IFN- β cDNA coding sequence was inserted as a 667-bp *BglII* fragment²³ into the single *BamHI* site of pSV532; a plasmid having the IFN- β gene inserted in the correct orientation with respect to SV40 late transcription was selected by restriction enzyme analysis and designated pSV532 cHFIF. Ap, ampicillin; Cm, chloramphenicol; Tc, tetracycline resistance markers. BB, *BamHI*-*BglII* hybrid site. The IFN- β mRNA transcription start site is indicated by an asterisk.



recombination in the recipient cell allows formation of a replicating SV40 derivative from which prokaryotic sequences have been removed. The gene for the major structural protein VP1 has been deleted (0.945–0.145 map units) from pSV529 and pSV532 and is replaced by unique restriction sites into which genes such as IFN- β can be inserted.

Monkey kidney (AP-8) cells were transfected with the resulting recombinant plasmids: pSV532 cHFIF, containing only the coding region of the IFN- β gene, and pSV532 gHFIF, containing in addition its promoter region (Fig. 1). Both constructions carry the IFN- β coding sequence in the correct sense with respect to SV40 late transcription, and also have the VP1 polyadenylation signal at their 3'-end. Thus, the formation of functional IFN- β mRNA is likely to occur. Note also that the interferon gene contains the signal sequence so that processing and secretion of the mature protein into the medium is expected. Indeed, up to 100 IU ml⁻¹ of human IFN- β could be detected. There was only a minor difference between pSV532 cHFIF and pSV532 gHFIF; this is remarkable as only pSV532 cHFIF contains the SV40 donor and acceptor sequences required for splicing of the precursor to the major late 16S mRNA. In pSV532 gHFIF, this region is replaced by the 286-base pair (bp) IFN- β 5'-flanking region.

On treatment with poly(rI) · poly(rC) (induction was at 11, 23, 35 or 47 h post-transfection), no induction effect was ob-

served when AP-8 cells were transfected with pSV532 cHFIF (Table 1). After long time intervals there was even a drop in IFN- β production, which might be due to a negative effect of the inducer on cellular metabolism. This was even more pronounced at later times (data not shown). In contrast to the results observed for pSV532 cHFIF, a 10–30-fold increase in

Table 1 Inducibility of cDNA and genomic DNA IFN- β clones

Hours post-infection	Interferon titre (IU ml ⁻¹)			
	11	23	35	47
pSV532 cHFIF				
– poly(rI) · poly(rC)	3	100	32	32
+ poly(rI) · poly(rC)	3	100	32	10
pSV532 gHFIF				
– poly(rI) · poly(rC)	3	32	32	10
+ poly(rI) · poly(rC)	6.3	1,000	316	32

AP-8 cells were transfected with pSV532 cHFIF or pSV532 gHFIF. At the indicated times after transfection, poly(rI) · poly(rC) (100 μ g ml⁻¹) was added for 1 h. The cells were then washed twice with phosphate-saline buffer and covered with fresh medium. Samples for interferon assay were taken 12 h post-induction. Control cells were treated in the same way, except that no poly(rI) · poly(rC) was added. Monkey interferon was not or was only barely (≤ 3 IU) detectable in the assay involving human FS-4 cells. The titres have been normalized on the basis of the NIH IFN- β reference G-023-902-527.

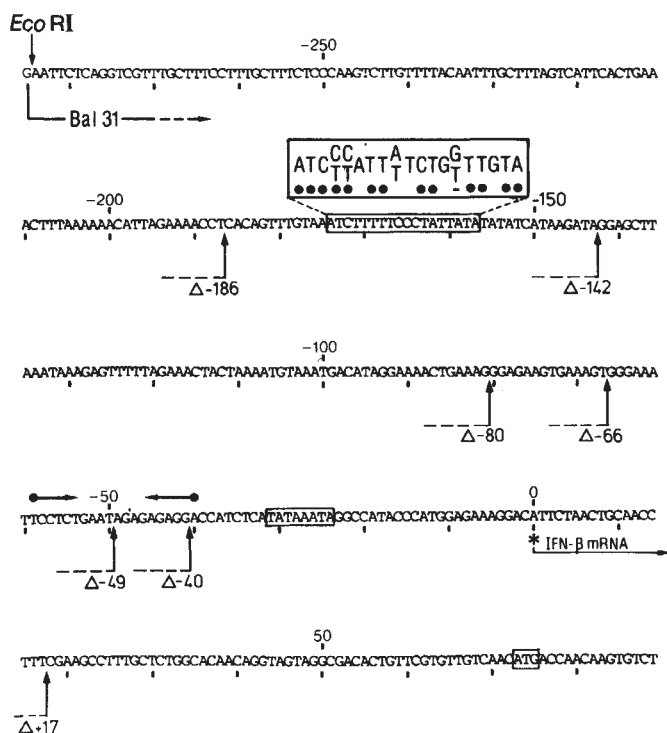


Fig. 2 Nucleotide sequence of deletion mutants in the human IFN- β upstream region. Only the non-coding strand is shown, with the TATA sequence and the ATG start codon enclosed in boxes. Nucleotide numbering starts from the presumed IFN- β mRNA capping site (indicated by an asterisk). All deletions originate at the *EcoRI* restriction site and extend to the upwards pointing arrows. They are indicated by the number of the nucleotide to which the deletions extend. The box at positions -174 to -157 indicates a sequence showing homology with a progesterone-receptor complex affinity site (ref. 2; large box above). Black dots indicate homology between both, a dash represents the absence of a nucleotide. A palindrome present around nucleotide -49 is also indicated.

interferon production was observed on poly(rI) · poly(rC) treatment of pSV532 gHFIF-transfected AP-8 cells. This effect was optimal for poly(rI) · poly(rC) induction at 24 h post-transfection. Little or no (≤ 3 IU) monkey interferon could be detected (Table 1). Analogous results were obtained when a 1.9-kilobase (kb) *EcoRI* fragment encompassing the coding region and its 5'- and 3'-flanking regions were inserted in the opposite orientation. In this case also, a constitutive production was evident, which again increased 10–30-fold on poly(rI) · poly(rC) treatment (data not shown). Other groups have also reported that the cloned 1.9-kb IFN- β gene responds to poly(rI) · poly(rC) induction in heterologous cells^{24–27}. Our results provide strong evidence for the role of the 5'-flanking region in the regulation of the human IFN- β gene.

In the next step, we generated deletion mutants by *Bal31* exonuclease treatment of the *EcoRI*-cleaved pg4ΔH plasmid vector (Fig. 1). After filling-in to ensure the formation of blunt ends, *EcoRI* synthetic linkers were added. Then, the *EcoRI*-*Bgl*/II fragments of the deletion mutants (carrying the IFN- β gene and its shortened 5'-flanking region) were transferred to *EcoRI*-*Bam*HI-cut pSV532, resulting in the pSV532 gHFIF-Δn series. Thus, the remaining IFN- β promoter regions were always ligated to the same SV40 sequence, minimizing structural and functional variability due to neighbouring sequences. After preliminary restriction mapping, the extent of the deletions was established by DNA sequencing. The nucleotide sequence of the deletion mutants is presented in Fig. 2, and their inducibility by poly(rI) · poly(rC) in Fig. 3. A sharp drop in inducibility was found when sequences were deleted between nucleotides -186 and -142. It is of interest that this area encompasses a sequence showing remarkable homology with a

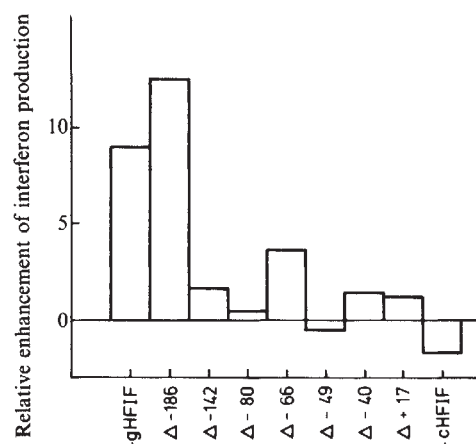


Fig. 3 Poly(rI) · poly(rC)-dependent inducibility of deletion mutants in the IFN- β promoter region. The figure shows the enhancement of human IFN- β production on poly(rI) · poly(rC) treatment (1 h, 100 $\mu\text{g ml}^{-1}$ at 24 h post-transfection; samples were taken 18 h post-induction) of AP-8 monkey kidney cells, transfected with a series of deletion mutants in the IFN- β promoter region. The values given are the average of three independent experiments each done in triplicate and assayed twice. In a typical experiment, ~ 30 IU ml^{-1} of human IFN- β , due to SV40 late promotion, were detected for all deletion mutants—this could be increased to ~ 300 IU ml^{-1} on induction.

consensus sequence found in the 5'-flanking region of steroid hormone responsive genes (for example, chicken oviduct proteins²), which might be involved in binding the progesterone-receptor complex. It is possible that there is an analogy between the mechanism of action of double-stranded RNA- and steroid hormone-dependent induction. The first part of the aforementioned motif can also be found in IFN- α genes arranged tandemly around position -213 (ref. 28). Unlike the IFN- β gene, IFN- α genes are induced only on viral infection. It is, however, not possible at present to exclude an additional role of the coding region in this induction process. Fusion of the 5' promoter region to other genes should clarify this.

Part of this work has been presented at a Royal Society meeting²⁹. This research was supported by grants from the Geconcerteerde Akties of the Belgian Ministry of Science and from the Belgian Fonds voor Geneeskundig Wetenschappelijk Onderzoek (no. 3.0063.81). We thank W. Burm for the interferon assays and W. Drijvers and S. Kaplan for help with the manuscript.

Received 12 October; accepted 10 December 1982.

1. Stewart, W. E. II *The Interferon System* (Springer, New York, 1979).
2. Mulvihill, E. R., Le Penec, J. & Chambon, P. *Cell* **24**, 621–632 (1982).
3. Nagata, S. *et al. Nature* **284**, 316–320 (1980).
4. Goeddel, D. V. *et al. Nature* **290**, 20–26 (1981).
5. Maeda, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7010–7013 (1980).
6. Derynck, R. *et al. Nature* **285**, 542–546 (1980).
7. Taniguchi, T., Fujii-Kuryama, Y. & Muramatsu, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4003–4006 (1980).
8. Goeddel, D. V., Shepard, H. M., Yelverton, E., Leung, D. & Crea, R. *Nucleic Acids Res.* **8**, 4057–4074 (1980).
9. Gray, P. W. *et al. Nature* **295**, 503–508 (1982).
10. Devos, R. *et al. Nucleic Acids Res.* **10**, 2487–2501 (1982).
11. Nagata, S., Mantei, N. & Weissmann, C. *Nature* **287**, 401–408 (1980).
12. Lawn, R. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 5435–5439 (1981).
13. Tavernier, J., Derynck, R. & Fiers, W. *Nucleic Acids Res.* **9**, 461–471 (1981).
14. Degraeve, W., Derynck, R., Tavernier, J., Haegeman, G. & Fiers, W. *Gene* **14**, 137–143 (1981).
15. Houghton, M. *et al. Nucleic Acids Res.* **9**, 247–266 (1981).
16. Ohno, S. & Taniguchi, T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5305–5309 (1981).
17. Gross, G. *et al. Nucleic Acids Res.* **9**, 2495–2507 (1981).
18. Lawn, R. M. *et al. Nucleic Acids Res.* **9**, 1045–1052 (1981).
19. Mory, Y. *et al. Eur. J. Biochem.* **120**, 197–202 (1981).
20. Gray, P. W. & Goeddel, D. V. *Nature* **298**, 859–863 (1982).
21. Taya, Y. *et al. EMBO J.* **8**, 953–958 (1982).
22. Gheysen, D. & Fiers, W. *J. molec. appl. Genet.* **1**, 385–394 (1982).
23. Derynck, R. *et al. Nature* **287**, 193–197 (1980).
24. Ohno, S. & Taniguchi, T. *Nucleic Acids Res.* **10**, 967–977 (1982).
25. Hauser, H. *et al. Nature* **297**, 650–654 (1982).
26. Zinn, K., Mellon, P., Ptashne, M. & Maniatis, T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4897–4901 (1982).
27. Canaan, D. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5166–5170 (1982).
28. Ullrich, A., Gray, A., Goeddel, D. V. & Dull, T. J. *J. molec. Biol.* **156**, 467–486 (1982).
29. Fiers, W. *et al. Phil. Trans. R. Soc. B* **299**, 29–38 (1982).

Working on the weather

Theodore K. Rabb

AN OLD joke has it that everyone talks about the weather, but nobody does anything about it. Perhaps with that in mind, Hubert Lamb has entitled the last chapter of his book "What can we do about it?". He is referring to forecasting, and to the policy decisions that will be necessary if mankind is to adjust in the best possible way to the changes climate may impose on us in the years ahead.

As the old joke might have led us to expect, however, this last quarter of Lamb's book, devoted to the present and the future, is the most tentative. Cycles of different lengths point in different directions; some long-term trends, such as the increasing concentrations of carbon dioxide in the atmosphere, foretell a gradual warming; others, such as the expected onset of another ice age, suggest the contrary; how the trends will balance each other out we do not know, and unforeseen occurrences, such as volcanic activity, may upset all predictions. In other words, despite the enormous progress of meteorology and the science of climatic reconstruction over the past few decades, we are still very far from knowing what will happen, or, as a consequence, what we ought to do about it.

That we have some sense of what the elements of a reasonable prediction might be, however, is a great advance. For centuries it has been assumed that climate is essentially unchanging; indeed, according to Lamb, that view has been undermined only since 1960. The result has been the transformation of climatology, and if in the process prediction has become far more complex, the achievement of greater understanding has nevertheless been notable. Above all, the past has become accessible to a degree previously unimagined — both in chronological span and in precision of detail. The creation of this climatic "laboratory" has been a signal accomplishment, inspiring scientific imagination in many fields and promising a firm basis for the ultimate goal of accurate prediction. Lamb has been one of the chief pioneers of this transformation, and it is fitting that he should have written a book that summarizes the state of the field after a quarter of a century of heady progress.

The range of the research of the past few decades is quite extraordinary. Apart from the traditional historical endeavours of finding and explaining written references to the weather, there have been dozens of highly ingenious explorations of physical data — tree rings, stalactites, sediments on

Climate, History and the Modern World. By H.H. Lamb. Pp.387. Hbk ISBN 0-416-33430-X; pbk ISBN 0-416-33440-7. (Methuen: 1982.) Hbk £16.95, \$33; pbk £8.95, \$16.95.

lake bottoms, polar ice and peat bogs — that have been used as proxies for temperature, wetness, wind direction and velocity, and storminess. All of those materials, built up in layers, can be subjected to techniques of reconstruction, such as carbon dating and other forms of isotope chemistry, which can recreate the meteorological conditions under which a particular tree ring, pollen sediment or segment of ice was laid down at a specified date. The result has been the creation of chronologies going back tens of thousands of years, and gaining ever finer resolutions as we approach the present. Thus Lamb can give us a weather map for the winter of 1431–1432, and daily maps for the Armada expedition of 1588.

The consequences of all this work are readily apparent in the first three-quarters of *Climate, History and the Modern World*, where, looking to the past rather than the future, Lamb is able to give a solid, richly nuanced account of the main features of the Earth's climatic history. He explains the operations of the physical system, the methods by which it has been investigated and the principal conclusions that have been reached for every distinct period from the last ice age to the present. In outline, the story is now fairly well established — a succession of warmings and coolings, each with its own distinct effects, some beneficial, some harmful. A period of gradual warming was apparently interrupted around 3000 BC, but by then the great deserts of Africa had been formed. Another major interruption, with glacier advance, came around 1500–1200 BC, when many settlements at high altitude were abandoned. The great ages of Greece and Rome were accompanied by a warm, fairly dry period, which gave way in the fifth century, as the Roman empire collapsed, to more unsettled, and then colder and wetter times. The later Middle Ages were again warm, but a significant break, starting in the thirteenth century, led to what has been called the "little" ice age. Except for brief interludes, such as the early sixteenth century, a colder and wetter regime held sway until the eighteenth century, when a warming trend began that continued almost to the present, certainly to around 1950. Similar, though not

exactly parallel, patterns can be seen in non-Western regions, though these have been less intensively studied.

Within these broad outlines, there was, of course, wide variation. The benign decade of the 1730s, for example, was followed by two harsh years, from 1739 to 1741. And the terrible decades of the 1430s and 1690s — easily the most severe in the period that is best documented, the last millennium — were perceived to be much worse even than the poor decades that preceded and followed them. It was precisely the increase in variability that was one of the worst features of difficult times, especially as they affected agriculture, health, communications and living conditions. For it is the latter set of considerations that is the ultimate concern of studies of climate. How is man affected? When is the climate favourable, and when not? The answer is not just a matter of temperature or wetness, because communities have flourished in widely diverse circumstances. In some cases, drying helps; in others, it harms. In general, it is stability, and thus predictability, that allows effective societies to develop and take root. The single worst consequence of a climatic regime is a demand for constant change and adaptation. At such times, the negative impact on human history, forcing withdrawals from settlements, trade routes, and occupations, has been unmistakable — the depopulation of large territories, from Africa to Greenland, the abandonment of the Great Silk Road and the voyages of the Chinese and the Polynesians, the setbacks of the late Middle Ages, and the rise and fall of such trades as mining, fishing and timber. Man-made alterations of the climate are likely to hold similar perils.

It is when he comes to positive results that Lamb is less convincing. Nobody can mistake the reaction to great droughts, storms or other disasters; nor can one fail to see that slower, but equally devastating shifts in habitability, such as a reduced growing season or the relocation of fishing grounds, can have immense historical consequences. But does a favourable climatic regime actually promote cultural and other achievements, or does it merely not impede them? This is not a distinction that Lamb recognizes, for he at least implies that a magnificent cathedral was built because it was in a rich agricultural community with a genial climate (p.51); that artistic creativity was stimulated (p.236) and kingdoms were united (p.277) as a result of weather patterns; and that the greatest advances in

health, technology and agriculture took place during the recovery from the "little" ice age after 1700 (p.308). For the historian such connections — like the attempt to link unsettled meteorological times with rebellion and revolt (p.196) — require much more precise chains of evidence and reasoning. How much store can one set by them if rebellions also occur in benign ages or if non-climatic influences seem quite sufficient to explain, say, the union of England with Scotland? A great deal more work will be needed before such arguments gain wide acceptance.

In the meantime, Lamb's remains the best introduction to the subject — erudite, wide-ranging and consistently thought-provoking. The book is not always as clearly written as it might be; the detail is sometimes overwhelming and contradictory; and there are repetitions and reversals of chronology (for example on p.237) which may confuse. But the call for further research, promising important discoveries about the nature of climate and its effects, is eloquently delivered. In particular, Lamb notes, it is essential to collate the historical research that has begun, and to shape current investigations into a coherent and unified effort. Already the studies of the past have revealed some remarkably consistent cycles and such extraordinary phenomena as the 0.75 correlation between the number of days with southwesterly surface wind in London

and the amount of snow deposited at the South Pole (p.270). Only if all such data are brought together, inconsistencies or contradictions resolved and accurate long time-series constructed, will we assemble a truly reliable account of past climates. Without this data bank — which, in current financial circumstances, remains a Utopian aim — it will be extremely difficult to document links with historical change, devise realistic policies or improve the

efficiency of forecasting. Lamb has given us a comprehensive picture of where we stand now; unless we take his advice, and move on to a new level of understanding, both of the past and therefore of the future, the old joke will remain true for generations to come. □

Theodore K. Rabb is a Professor of History at Princeton University. He is co-editor of Climate and History (Princeton University Press, 1982).

The ins and outs of sexual reproduction

Mark Ridley

Philopatry, Inbreeding, and the Evolution of Sex. By William M. Shields. Pp.245. Hbk ISBN 0-87395-617-6; pbk ISBN 0-87395-618-4. (State University of New York Press: 1982.) Hbk \$49, £39.20; pbk \$16.95, £13.55.

Philopatry, Inbreeding, and the Evolution of Sex is the latest answer to the question of why sexual reproduction exists. It is a long argument, drawn from many parts of evolutionary biology: Shields calls it a "teleonomic web", but, in the metaphor of zoological artefacts, it is more a bird's nest than a spider's web. He has selected sticks from the literature of population genetics, speciation, life histories and behaviour, added a few feathers of his own, and arranged them according to a most personal design.

Shields starts with philopatry (which means not moving far from place of birth to place of breeding). A philopatric species may be broken into many local, relatively inbred populations; this Shields turns round to say that inbreeding is the function of philopatry. He then examines the advantages and disadvantages of inbreeding, accepts the former, rejects the latter, and concludes that organisms inbred much more than biologists have supposed.

The puzzle of sex is that asexual reproduction has two clear-cut advantages: it saves on males, and it propagates genes more efficiently. If a (dominant) mutant arises, causing its bearer to produce a proportion of its offspring asexually, the mutant will go into 50 per cent of the asexual offspring (for the gene is heterozygous when rare), compared with 25 per cent of its sexual offspring (after crossing with a wild-type homozygote). But those sexual offspring are now infected, and the sexual part of the population will gradually be taken over as further mutants, increasing the proportion of asexual offspring, arise and spread.

Such is the disadvantage of sex, but it assumes outbreeding. If there is inbreeding sex costs less. Under inbreeding (as W. D. Hamilton first noticed) sex ratios may be female-biased, which saves on males. In-

breeding also protects a sexual population from genes for parthenogenesis: in the extreme case of self-fertilization, sexual reproduction becomes equivalent to asexual.

At this point, Shields could have stopped. It is widely accepted that sex is less costly with inbreeding, but not that inbreeding is common, so inbreeding would be an original answer to the question. But he goes further, and argues that sex with inbreeding is not just equivalent to parthenogenesis, but better than it. He draws on an argument which has come to be called Muller's ratchet. Here (as elsewhere) the case is inconclusive, not least as we find Shields confidently unaware that Muller's ratchet provides only a long-term, group selectionist advantage for sex.

Shields's main thesis, however, is that inbreeding is much more common than usually supposed. Biologists have been persuaded that inbreeding is exceptional by many observations of avoidance of inbreeding, by the manifestation of inbreeding depression (together with an explanation of why it exists) and by evidence that sex ratios are not usually female-biased. Shields does indeed take on all of these arguments, but he is not really stimulating. For instance, he denies the evidence for inbreeding depression because it has mainly come from laboratories. He also argues that inbreeding is advantageous because it speeds up the rate of evolution (which is group selectionist) and protects coadapted gene complexes (which are imaginary).

Although I think Shields is wrong, I would still recommend his book to its intended audience. It is confined, by its language (which reads like, say, the *American Naturalist*) to experts. They will find in it an original attempt to solve an important and unanswered question. And, even if the overall theory is wrong, perhaps someone will find a use for one or two of those feathers which Shields himself has contributed. □

Mark Ridley is the Hayward Junior Research Fellow of Oriel College, Oxford University.

ADVERTISEMENT

The ANS Journals

Your number one source for current information on the nuclear industry.

- NUCLEAR NEWS (NN)
- NUCLEAR SCIENCE & ENGINEERING (NS&E)
- NUCLEAR TECHNOLOGY (NT)
- NUCLEAR TECHNOLOGY/FUSION (NT/F)
- NUCLEAR STANDARDS NEWS (NSN)
- NUCLEAR REPORT (NR)
- TRANSACTIONS (TR)
- REMOTE SYSTEMS TECHNOLOGY (RSTD)
- ANSPI (ANSPI)

(Write for our complete publications catalog.)

Send to: American Nuclear Society
555 North Kensington Avenue
LaGrange Park, IL 60525 USA

Please send me information on...

- ☐ NN ☐ NS&E ☐ NT ☐ NT/F ☐ NR
☐ NSN ☐ TR ☐ RSTD ☐ ANSPI
☐ National Meetings ☐ Topical Meetings
☐ Individual Membership ☐ Corporation Membership
☐ Send me your free catalog

NAME _____

TITLE _____

ORGANIZATION _____

ADDRESS _____

CITY _____

STATE _____ ZIP _____

Knowing the need for the need to know

Niles Eldredge

The National Museum of Natural History. By Philip Kopper. Pp.493. ISBN 0-8109-1359-3. (Harry N. Abrams: 1982.) \$60, £38.50.

A FEW years ago, when the United States Congress was debating the merits of supporting archaeological salvage work in the valley of the Nubian Nile, a Member of the House of Representatives scornfully suggested that the citizens of the United States had better ways to spend their money than squandering it on the "paleontology of Egypt". Yet Congress has duly undertaken the yearly support of the Smithsonian Institution — a major complex of museums and research stations — ever since the Englishman James Smithson bequeathed his fortune "to found at Washington, under the name Smithsonian Institution, an Establishment for the increase & diffusion of knowledge among men". Among the array of organizations that have sprung from Smithson's original largesse, Philip Kopper has focused on one: the National Museum of Natural History. He has produced a verbal portrait that, on the whole, nicely augments the quite stunning parade of colour photographs that both dress this book and themselves have much to say about this particular museum and others of its ilk.

Why, indeed, do we need to know about the "paleontology" of Egypt? For it is this need to know that supplies the ultimate rationale underlying the existence of any natural history museum. There are, of course, other rationalia: some objects are precious, some are beautiful. Some are both: the Hope diamond is one of the National Museum's outstanding treasures. Archaeological pots and Indian bead work have soared in value in recent years — and their beauty, fragility and rarity must all be taken into account. Then, too, museums disseminate information to the world at large, and they *entertain*: many if not most daily visitors casually peruse an exhibit or two for mere diversion and simple pleasure. Kopper structures his book to embrace all of these diverse functions: the National Museum is a museum of "collections, conservation, science and art" as well as a "public museum". But it is first and foremost a museum of science, as are all its major sister institutions the world over.

Kopper's description of the scientific activities of the National Museum — the final one-third or so of the book — is of necessity rather light, but also somewhat disappointing. In the century following the

founding of the major natural history museums, we have seen the disappearance of cultures, species and many special archaeological and palaeontological sites. Conservation has only recently become all the rage (depressed somewhat by current economic conditions), and it is a blessing indeed that we have had far-sighted collectors filling these museums for so long. But we have also witnessed a return to more holistic (and even hierarchical) approaches to various branches of science. No longer, for example, are molecules the sole objects deemed worthy of serious enquiry in biology as they more nearly were only a decade ago. The contributions of museum scientists to theories of plate tectonics, systematics, biogeography and evolution emphasizes the importance museums have in the scientific world. Kopper, I feel, sees the National Museum too much in isolation; he rightly discusses the achievements of its staff, but robs them of the grander context of the scientific side of the entire academic world.

Yes, natural history museums are treasure troves and libraries of the physical objects of this and other planets. Kopper's book takes a good look at that side of such museums. But the "need to know" transcends curiosity, conservation and the quest for practical application: the urge to understand all aspects of the surrounding world is what has taken us to our present state, for better or worse. We can hardly afford to drop the enterprise now — as, fortunately, a majority of Congressmen seem to perceive. □

*Niles Eldredge is at the American Museum of Natural History, New York. His most recent book, written with Ian Tattersall, is *Myths of Human Evolution* (Columbia University Press, 1982).*

Nervous disposition

Jonathan Cooke

Neuronal Development. Edited by Nicholas C. Spitzer. Pp.424. ISBN 0-306-40956-9. (Plenum: 1982.) £31.50, \$45.

RECENT decades have seen elucidation of the general strategy of cellular events whereby nervous systems are constructed. Within this framework the search for detailed mechanisms must be made and vast numbers of laboratories now produce parallel stories, using various model systems believed to render one or other part of the developmental sequence especially transparent. Any multi-author work that samples the field, however carefully assembled, has therefore a patchiness

about it. It may give us the flavour of current investigations, but is unlikely to serve as a definitive reference work. The present book is no exception.

The first phase of CNS development remains the most mysterious. Accordingly, only two of the contributors claim to address this process in which a spatial pattern of small cell groups, each programmed to give rise solely to one piece of the final structure, arises by position-dependent determinations within one layer of the embryo. One of these chapters is a definitive review but describes an atypical organism, the leech, which utilizes defined cell lineage even for founding the pattern. The other is a *tour de force* of wilful iconoclasm, claiming to dismiss the established tradition of evidence that the vertebrate CNS pattern arises by position-dependent determinations among a sheet of some 5,000 ectoderm cells in the gastrula, and to implant instead the idea that it is founded from seven small groups of determined precursors set up at early cleavage stages. It certainly makes the critical researcher think, but the presentation is insufficiently balanced or authenticated for a book of this type. The polyclone concept is attractive apart from its truth value, however, and neurobiologists unversed in vertebrate experimental embryology may well, as it were, buy it.

The grasshopper embryo offers a beautiful system in which to study the next phase of development, where detailed local assemblies of neuroblasts are set up, probably by lineage mechanisms. This is reviewed in awesome but exquisite detail by Corey Goodman. A further excellent chapter by Palka lays out the complex but almost coherent story that somatic genetic analysis is revealing, about how the organization of insect ectoderm development by "compartments" is reflected in and utilized by the repetitive pattern of the CNS.

There is much else of interest, about neurite behaviour and the reliable guidance of fibre tracts towards their general target areas, but it is in respect of perhaps the final phase of the strategy — that whereby specific patterns of connections *within* target areas are assured — that another imbalance in the book occurs. Only the last chapter describes a system known to show such neuronal specificity, among incoming optic fibres to the vertebrate tectum, and here this is treated almost incidentally in what is otherwise a specialized report on anatomical regeneration.

In short, the book does indeed convey a fairly coherent view of what is happening in research into neuronal development, and it includes some fine reviews. But a certain arbitrariness and some omissions of representation are obtrusive. □

Jonathan Cooke is a member of the Scientific Staff at the National Institute for Medical Research, Mill Hill, London.

● Paul Davies's *The Edge of Infinity: Beyond the Black Hole* (reviewed in *Nature* 294, 488; 1981) is published today in paperback by Oxford University Press. Price is £2.95.

The official's book of risky business

Alastair Hay

Major Technological Risk: An Assessment of Industrial Disasters. By Patrick Lagadec. Pp.536. ISBN 0-08-028913-4. (Pergamon: 1982.) \$60, £30.

ONE way of summarizing the message of Paul Lagadec's book would be "High technology versus democracy". The result of a doctoral thesis in political science, Lagadec's *Major Technological Risk* is a panoramic look at the problems posed by high technology. As befits a thesis on this subject, it is the politics of risk which interest Lagadec rather than the idea of finding a "technological fix" for every conceivable accident.

The answer to many of the problems involved in assessing risk is information, says Lagadec. And the most important element in this is informing the public. Risk assessment is about probabilities and uncertainties. These ideas, he points out, may be understandable to the expert; they are not so to the average citizen who is often running the risk. Lagadec believes that the public wants this information, but that industry and governments may not always be too keen to provide it.

To illustrate his theme the author chooses five seminal accidents of the last decade: the chemical explosion at Flixborough (United Kingdom); the release of dioxin at Seveso (Italy); the failure of the nuclear reactor at Three Mile Island (United States); the discharge of oil from the wrecked tanker Amoco Cadiz (France); and the release of chlorine from tanks on a derailed train in the centre of Toronto (Canada). In all of these cases technical failure was the direct cause of the accident, although the post-mortems of the events showed that inadequate legislation was a contributing factor in some, and in others communication between emergency services was so poor that this had its effect on the outcome.

It is the aftermath of these accidents which most concerns Lagadec. Mapping the extent of the damage was relatively easy in the case of Flixborough and the Amoco Cadiz; the trail of destruction was evident from the blast-damaged buildings and oil-polluted beaches. Apportioning blame in both cases was less easy, and Lagadec argues that the official Commission of Enquiry into the Flixborough accident was too lenient in this area.

The cases of Seveso and Three Mile Island illustrate points different from those raised by Flixborough. Assessing the damage was much more difficult; the danger was not visible. At Seveso it took time to find the extent of dioxin pollution — the delay in doing this resulted in people living in a highly polluted area for two weeks, before they were told to leave their homes. At Three Mile Island the release of radiation caused panic amongst the local

populace, and an estimated 200,000 people left the area of their own volition. They were encouraged to take this step by the ineptitude of the government, regulatory authorities and company officials who issued often contradictory information.

At Seveso, poor communication between officials and the residents rapidly led to disillusionment with the health surveillance programme. The public began to distrust officials. Similar sentiments were evident in the wake of the Three Mile Island debacle. According to Lagadec the consequences of this public distrust are serious. If it persists, he argues, it will become even more difficult to contain disasters. His remedy is for officialdom to take the public into its confidence. By so doing the average citizen may at least be convinced that everything possible is being done to ameliorate the effects of the disaster.

As for the release of chlorine in Toronto, this is an example of officials and emergency services working well. The evacuation of over 250,000 people was achieved relatively painlessly. There are lessons here for regulatory authorities, Lagadec says.

This book is addressed to those officials responsible for dealing with potential dis-

aster. It might also be aimed at an informed public — the text, if somewhat diffuse, is not too difficult for the lay reader. As for workers in industry, they do not fit into Lagadec's scheme of things. He argues that they will benefit indirectly from more public accountability. Pollution controls such as Clean Air Acts do, of course, benefit everybody but workers in industry have their own particular problems, made all the more difficult by inadequate information — their requirements are more direct than the prescriptions offered in this book.

The book does suffer from being too long and is in parts repetitive. The index — a list of accidents — is inadequate for the text, and led to considerable frustration on the part of this reader. More specifically, with regard to the cause of the Seveso accident the author is mistaken in saying that this was probably attributable to a reduction in the solvent in the 2,4,5-trichlorophenol reactor; localized heating in the top few centimetres of the reaction mix was almost certainly to blame. These criticisms, however, do not detract from the value of this book. Its message is important, and those interested in risk would do well to read it. □

Alastair Hay is a Lecturer in the Department of Chemical Pathology at the University of Leeds, and author of The Chemical Scythe: Lessons of 2,4,5-T and Dioxin, published late last year by Plenum Press.

Art of gel cookery

E.G. Richards

Gel Electrophoresis of Nucleic Acids: A Practical Approach. Edited by D. Rickwood and B.D. Hames. Pp.242. Pbk ISBN 0-904-14724-X. (IRL Press: 1982.) £8.50, \$18.

LIKE a *soufflé* dish, gel electrophoresis requires no great knowledge of theory to achieve admirable results but it does require a certain concentrated attention to practical detail. This volume provides such detail in abundance for the application of gel electrophoresis to nucleic acids. These details are described by ten authors in six chapters which cover more ground than the title of the book might suggest, including as they do, articles on the sequencing of both DNA and RNA and on the electrophoresis of nucleoproteins. Detailed descriptions of apparatus and protocols for their use are provided for the analysis of RNA or DNA preparations in agarose and polyacrylamide gels, gradient gels, two dimensional gels, preparative gels and most of the other available methods that are applicable to nucleic acids.

Theory, even at the level of the application of Ohm's law, is virtually absent. One consequence of this is that the naive reader has no way of knowing whether he

will stumble into a pitfall if he departs ever so slightly from the received instructions. He may even conclude that all procedures have been carefully planned to give optimal results. But every cook has his favourite recipes and several of the authors admit that theirs was arrived at by a process of trial and error. Even if the reader sticks to the letter of the law strange results will sometimes appear as disaster strikes. The explanations of some of the more controllable of these are described; others which still defy explanation are not mentioned.

The scanning of gels to obtain an estimate of the amount of nucleic acid in a zone is mentioned several times. It is a pity that an additional chapter was not included which described in detail how this might be accomplished and some of the attendant difficulties, for the optics used in some commercially available gel scanners are barely equal to the task.

Workers interested in the theory of gel electrophoresis or in the history of its application to nucleic acids had better look elsewhere, but this laboratory manual will be found useful if you want to set up experiments. The list of worldwide firms which supply equipment and materials, included at the end, will help you too. □

E.G. Richards is a Senior Lecturer in the Department of Biophysics, King's College, University of London.

Patterns of life in the sea . . .

J.H.S. Blaxter

Ocean Management. Vol. V, Pt 1 of *Marine Ecology: A Comprehensive, Integrated Treatise on Life in Oceans and Coastal Waters*. Edited by Otto Kinne. Pp.642. ISBN 0-471-27997-8. (Wiley: 1982.) £39.50, \$87.50.

PROFESSOR Otto Kinne here presents us with first volume of a final trilogy in his long series *Marine Ecology*, started in 1970. The overall title for the final three volumes is *Ocean Management* but this first one, which covers zonation and organismic assemblages, cannot be said to contain anything about the practice of management.

In his introduction, Professor Kinne writes of the intention to show how organisms act in concert at the supra-individual or multi-species level. He then hands over to Professor J.M. Peres who has written the remaining eight, substantial chapters in the book. Professor Peres has shown great versatility in bringing together a remarkable range of information which he handles in a readable, attractive style. This is essentially a review of work on general concepts and summaries of biogeography and trophic relationships, the object being to find an underlying pattern within the pelagic and benthic zones and in the interrelations between the various zones. The reviewing is particularly good on the French and Russian literature — indeed this type of approach to marine biology is a particular strength of the Russians.

Among the varied ecological niches described in the book, one may pick out some which have been investigated in more recent years. The sulphide system or thiobios is the reduced sediment layer, with a high content of hydrogen sulphide, in which a surprisingly wide range of organisms live, including metazoans. The phreatic fauna inhabit coastal subsoil systems where fresh water passes under the beach and progressively mixes with sea water. Here protozoans, mainly ciliates, turbellarians and oligochaetes may be numerous. In the pelagic zone the fauna associated with floating objects such as large animals, flotsam, *Sargassum* weed or lumps of oil, is very varied; it may be epiphytic or only seeking shelter.

Little is said, however, about the work on frontal systems. These are regions where waters of different types impinge, which may prove to be especially important areas for organisms to concentrate, giving, for example, good feeding for the delicate and vulnerable larval stages of fish and invertebrates. The treatment of vertical migration is also disappointing. The subject is not even mentioned in the subject index, yet it is such a widespread phenomenon — both geographically and biologically — that a full discussion of its biological role and

importance would have been apposite. Nevertheless, the importance of the sinking of food material from the surface in regions of high productivity, initiating the so-called "ladder of coprophagy", is discussed and thoroughly so.

The painstaking nature of the book is revealed by the author, species and subject index, which occupies the last 60 pages, and by the extensive bibliographies to be found at the end of each chapter. But it may be that, in trying to be utterly comprehensive,

this series as a whole falls into a trap. Assuming the final volume appears some time next year, the series will have been brought out over a period of 13–14 years and the earlier volumes will be very dated. Professor Kinne and the publishers should be encouraged to publish some form of year book, as produced by *Encyclopedia Britannica*, which will help to maintain *Marine Ecology* as a much-used and valuable source of reference. □

J.H.S. Blaxter is Senior Principal Scientific Officer at the Dunstaffnage Marine Research Laboratory of the Scottish Marine Biological Association, Oban, Scotland.

. . . and of inequality before death

W.F. Bynum

Death is a Social Disease: Public Health and Political Economy in Early Industrial France. By William Coleman. Pp.322. ISBN 0-299-08950-9. (The University of Wisconsin Press: 1982.) \$35, £28.

IN THE beginning was the Word, but now there is mostly the Number. We live by the clock, judge individual success in terms of salary, and collective success (or failure) by industrial output and inflation rate. We count heads regularly, keep an eye on mortality, morbidity and fertility levels, and we expect our scientists to produce statistically significant results.

We in the United Kingdom also tend to tell the story of this development in British terms, in a path leading from the seventeenth-century political arithmetic of John Graunt and William Petty, through the economic theories of Mill and Ricardo, the demographic insights of Thomas Malthus, the hedonistic calculus of Bentham, the social and vital statistics of Edwin Chadwick and William Farr, to the more sophisticated mathematics of Francis Galton and Karl Pearson. The great divide is the Industrial Revolution, based on the factory system, coinciding with the modern rise in population, and accentuating the process of urbanization. Because England was the first industrial nation, the introspective preoccupation is understandable, and both historians of medicine and of the social sciences have noted the relationship between political economy and public health in early Victorian England. This can be seen most clearly in the career of Chadwick, the Benthamite lawyer who dominated the Poor Law Commission of 1832, became Secretary of the Commissioners created by the New Poor Law of 1834, and by the early 1840s was the acknowledged leader of Britain's public health movement. Chadwick's *Report on the Sanitary Condition of the Labouring Population of Great Britain* (1842) is justly seen as a classic liberal statement of the crucial importance of good health for the

economic and social well-being of the nation.

Two years earlier, in 1840, Louis René Villermé (1782–1863) published a massive study which shared some of Chadwick's assumptions, methods and conclusions. But Villermé's referent was France, and while his *Tableau de l'État Physique et Moral des Ouvriers employés dans les Manufactures de Coton, de Laine et de Soie* was instantly recognized in his own country as a major piece of sociomedical analysis, both the work and its author have remained too little appreciated this side of the Channel. Unlike Bichat, Corvisart, Laennec, Louis and the other leaders of French medicine during the period, Villermé was never translated, perhaps because his monographs and articles were deemed to be less internationally relevant than the pathological and clinical works of his peers. The relative historical neglect of Villermé has been partially remedied by several articles by French economic and social historians and the papers by the medical historian Erwin H. Ackerknecht. Now we have a major full-length study by William Coleman, a distinguished American scholar of French science and medicine.

Although Villermé is the principal focus of Coleman's book, this is not a conventional biography. The basic facts of Villermé's early life, medical education, surgical career in the army during the Napoleonic Wars and subsequent activity as a Parisian investigator of public health and adjacent issues are dispensed with in a dozen pages. Rather, Coleman uses Villermé's work as a vehicle for exploring the delicately related themes of industrial development, economic thought and sociomedical investigation in France during the first half of the nineteenth century. Coleman is equally interested in *réalité* and *mentalité*, in social conditions and in the ideological basis of French debates on that English import, the factory system. He thereby exposes more thoroughly than any previous historian the

economic background to modern public health and social medicine.

Industrialization gained much more momentum in France during the Restoration, although major French-speaking political economists such as Jean Baptiste Say and Simonde de Sismondi had lived through the Revolutionary period as adults. They shared with Villermé a passion for collecting and analysing social facts of a statistical kind, and had a liberal faith in the neutrality of their data. They were concerned above all with the creation and distribution of wealth, issues to which machine power and a rising entrepreneurial class lent particular urgency. Within the new industrial order, relations between employer and employee were magnified, partly because of the sheer number of workers which successful entrepreneurs could hire, partly because trade cycles and frequent bankruptcies created problems of unemployment. Increasingly Villermé's work forced him to come to grips with much more than simple mortality statistics.

He began with the army and ended with society at large. In between, however, came a series of studies on prisons: their social functions, economics and the health problems they created. Recent works by Michel Foucault, Michael Ignatieff and David Rothman have alerted us to the radical change in the early nineteenth-century philosophy of prisons. No longer to be simply coercive or custodial, penitentiaries were to be institutions of reform, instilling in their inmates a new sense of social responsibility and re-training them to take their place in society. The reality was often different, and Villermé's researches revealed appalling conditions and staggering mortalities in many Paris prisons. His book, published in 1820, alerted those who would take notice to the horrors of life behind locked doors. It also gave Villermé his first extended taste of statistical investigation. By the 1820s, however, the whole of Paris had become his oyster, and for almost two decades he busied himself with two related, seemingly simple questions: Why did people in cities die sooner, on average, than people in the countryside? Why did poor people, on average, die earlier than rich ones?

The answers Villermé provided changed a bit over the years as he sharpened his analytical tools and refined his data. The cholera epidemic of 1832 gave him a particularly striking opportunity to examine differential mortality ("Inequality before death", Coleman calls it) and he came increasingly to doubt the simple environmental explanations which eighteenth-century doctors had proposed. More fundamental, he discovered, was the brute fact of income: not simply where you lived, but how much money you had to pay for food, clothing and shelter. It may seem a banal conclusion, but it had polemical consequences of considerable magnitude in the 1830s.

Villermé's social diagnosis, however,

was bolder than his therapeutics. He campaigned hard for legislation protecting children from industrial exploitation, but, for the rest, social melioration would come not through changes in the ownership of the means of production, but through individual change of heart. Workers must learn sobriety and reliability, masters must cultivate benevolence and philanthropy. He accepted the existing social order, rather than, like his contemporaries Marx and Engels, wishing to change it.

These and other themes are lucidly discussed in Coleman's book, which reminds us that concern over public health was not a British monopoly and examines in illuminating detail the ideological dimensions of social statistics and social medicine. It is a study of the origins of our modern, numerical society. At a time when the philosophy of the welfare state is being challenged, it also has a poignant topicality. □

W.F. Bynum is at the Wellcome Institute for the History of Medicine, London.

Redouté's lilies



THE illustration — of Crinum jagus, an African amaryllid — is taken from Lilies and Related Flowers, recently published by Overlook Press in the United States and Michael Joseph in Britain. The book contains reproductions of about a hundred plates from nearly five hundred first published from 1802 to 1816 in the eight large volumes of Pierre-Joseph Redouté's Les Liliacées, a glorious series which has been oddly overshadowed by the fame of his paintings of roses. The plates are printed in colour, about half the size of the originals, and are accompanied by notes on the history, horticulture and taxonomy of the plants figured, contributed by Brian Mathew. A short account of Redouté's career and a bibliographical note on the original edition complete an attractive book. Price is \$60, £19.95. Sandra Raphael

Re-scattering theory

Colin Wilkin

Scattering Theory of Waves and Particles, 2nd Edn. By Roger G. Newton. Pp.740. ISBN 0-387-10950-1. (Springer-Verlag: 1982.) DM98, \$43.60.

A GENERATION of physicists studied quantum scattering theory from Mott and Massey's *The Theory of Atomic Collisions* (first published by Oxford University Press, in 1933), but in the mid-1960s two very useful books appeared which largely took over this role. Of the two, *Collision Theory* by Goldberger and Watson (Wiley, 1964) was the more physical whereas Roger Newton stressed the formal aspects of the subject and in particular the connection with classical scattering of particles and electromagnetic waves.

Many important papers have been written since then and in the new edition Newton has added a thousand new references. I must, however, confess to being rather disappointed that he has modified the text so little in response to this more recent work. To quote but one example, much of the prolific data on medium energy proton-nucleus scattering has been analysed within the framework of the eikonal approximation. The corrections to this approach and the relationship to the WKB method have been elucidated in a series of remarkable papers by Wallace but the only mention of them is to be found in the bibliography.

Major changes have been made in the sections on the three-body problem, where the original treatment was insufficient in view of modern developments, and on the three-dimensional analysis of potential scattering. The discussion on the inverse scattering problem, where one attempts to construct the potential from physical input such as phase shifts, has been much extended though it cannot compete with the specialist monograph such as Chadan and Sabatier's *Inverse Problems in Quantum Scattering Theory* (Springer-Verlag, 1977) which contains a long historical introduction by Newton!

Perhaps I am expecting far too much from a single book: even in the first edition Newton eschewed the suggestion that it should be considered as an encyclopaedia of scattering theory and this is even less feasible now. There are many attractive parts to the work such as the comparison of the treatment of rainbow and glory scattering of light, classical and quantum particles which is so illuminating for heavy-ion scattering. And then there are all those invaluable references. All in all I expect that it will remain a source-book where even experts will go to see things done "properly". □

Colin Wilkin, Reader in Physics at University College London, is Honorary Editor of Journal of Physics G: Nuclear Physics.